PETRE MUREȘAN

METODE MATEMATICE

ÎN CLINICĂ, LABORATOARE ȘI OCROTIREA SĂNĂTĂȚII

EDITURA MEDICALĂ

Dr. MUREȘAN PETRE

METODE MATEMATICE
IN CLINICĂ, LABORATOARE ȘI OCROTIREA SĂNĂTĂȚII

EDITURA MEDICALĂ
BUCUREȘTI
1963

Riduceo la "

se introduce noțiunea de probă, de apoi

pe REPEAT și apoi pe X.

CAPITOLUL I

NOȚIUNI GENERALE

1. INTRODUCERE

Statistica matematică este utilizată de mult timp în domeniul fizicii, chimiei, astronomiei. Utilitatea ei ca metodă de cercetare a fost recunoscută și de către cercetătorii din domeniul științelor naturale, în biologie și medicină.

Unii dintre teoreticienii calculelor probabilității au exagerat chiar în acest domeniu, arătînd că la baza cunoașterii în domeniul tuturor științelor nu stă certitudinea, ci probabilitatea.

Cei mai mulți dintre teoreticienii burghezi susțin că la baza acesteia stă legea numerelor mari, că numai teoria probabilității ne poate apropia de adevăr, adevăr prin care s-ar înțelege „caracterele medii" ale unor colectivități de fenomene, ființe sau obiecte, observate în masă.

Realitatea obiectivă n-ar exista, deci menirea științei ar fi doar, ca pe baza probabilității, să descopere aceste „caractere medii", care ar caracteriza, într-o măsură oarecare, fenomenele studiate.

Laplace, de exemplu, în lucrarea sa „Théorie analytique des probabilités", afirmă că toate cunoștințele noastre nu sînt decît probabile. Chiar și puținele cunoștințe, pe care le putem ști cu certitudine, de exemplu în matematică, sînt bazate pe probabilități.

Cu unele deosebiri, în ultimul timp, teoreticienii burghezi ai statisticii matematice exprimă, de fapt, și ei cele arătate anterior, exagerîndu-le chiar. Legea numerelor mari devine, astfel, dintr-o metodă, scop în sine, căutîndu-se să se renunțe la analiza calitativă a fenomenelor cercetate.

Pe baza acestei legi, se trag concluzii în ceea ce privește fenomenele, indiferent de esența materială, specifică fenomenului cercetat, iar problema tendințelor acestora se rezolvă pe baza formulelor matematice și nu pe baza studiului complex al cauzelor determinante și adjuvante.

★

Nu înseamnă însă că, o dată cu combaterea acestor concepții, ce fetișizează legea numerelor mari, să îndepărtăm, și partea ei pozitivă, și aplicabilitatea ei ca metodă importantă de cercetare, în toate domeniile științelor concrete, deci și în biologie și medicină.

Legea numerelor mari poate fi şi trebuie să fie utilizată ca metodă în toate cercetările experimentale şi mai ales în observaţiile fenomenelor sau colectivităţilor, care sub influenţa unor acţiuni multilaterale sînt supuse unor variabilităţi, mai mult sau mai puţin importante.

Acesta este cazul şi în biologie, îndeosebi în medicina clinică. Observaţiile clinice duc la rezultate variate, tocmai prin faptul că organismul uman sănătos şi mai ales cel bolnav prezintă o gamă întreagă de intensităţi în ceea ce priveşte pulsul, respiraţia, tensiunea, metabolismul bazal, tabloul sanguin, concentraţia diferitelor substanţe din sînge, lichid cefalorahidian, urină etc.

Într-o serie de boli, chiar şi în cele cu evoluţie în general clasică (tipică), există o mare diversitate în ceea ce priveşte reactivitatea organismului la agenţii patogeni şi factorii de mediu. Am putea spune că fiecare individ îşi are formele specifice, proprii de boală.

Aceasta nu înseamnă că nu ar fi asemănări importante în cadrul diferitelor boli şi că, deci, nu am putea cuprinde bolnavii în diferite entităţi nosologice.

Faptul este posibil, date fiind asemănările esenţiale în evoluţia unor boli la colectivităţi de indivizi, chiar cu reactivităţi diferite ale organismului. Mai mult, clasificarea este necesară, deoarece numai prin observaţii pe colectivităţi putem trage concluzii generale valabile. Or, scopul practic al oricărei ştiinţe este tocmai generalizarea concluziilor ce rezultă din experiment sau observaţie.

★

În fenomenele atît de variate, cum sînt cele din cadrul medicinei, calculul probabilităţilor devine o metodă de mare valoare în sinteza observaţiilor clinice sau a experimentelor de laborator, metodă necesară concluziilor juste, pe care le putem generaliza.

Cu toate acestea, statistica matematică, deşi cunoscută şi utilizată în trecut în medicină de o serie de cercetători, nu a reuşit, totuşi, pînă în prezent, să se afirme ca o metodă importantă de cercetare.

În ultimul timp, s-a reuşit a se face cunoscută importanţa acesteia în studiul stării de sănătate a populaţiei : demografie, morbiditate, statistica activităţii medico-sanitare şi în special în studiul dezvoltării fizice a populaţiei.

Parcurgînd publicaţiile medicale sau tratatele medicale se constată că acestea conţin multe date statistice. Sînt suficiente exemple, în care se poate constata că utilizarea metodei statistice în cercetările şi concluziile respective a fost judicios făcută.

Există însă cazuri în care se poate constata că autorii nu au utilizat în mod just metodologia statistică, putîndu-se, din această cauză, prejudicia materialul cules (uneori cu multă migală şi cheltuieli), cît şi concluziile cercetării.

Cît de mare trebuie să fie colectivitatea cercetată sau de cîte ori trebuie repetat experimentul respectiv pentru a putea trage concluzii juste, este o problemă puţin cunoscută de către unii cercetători. Statistica matematică poate ajuta pe cercetător în stabilirea mărimii colectivităţii sau a numărului de experimente necesare.

Metodele matematice, în cercetările clinice şi experimentale, ca şi în cercetările de teren, trebuie considerate, deci ca metode deosebit de

importante. Deşi nu pot înlocui analiza calitativă a esenţei materiale a fenomenelor sau a colectivităţilor cercetate, ajută, totuşi, într-o măsură importantă, la concretizarea şi sintetizarea concluziilor rezultate din cercetările efectuate, deci constituie un factor de bază în generalizarea lor, în progresul ştiinţei. Aceasta, cu atît mai mult, cu cît în biologie şi în special în medicina clinică, variabilitatea mare a fenomenelor este o regulă aproape generală.

Fără utilizarea metodelor matematice nu putem fi, de cele mai multe ori, siguri de existenţa unor erori, care pot fi atît de importante, încît să falsifice rezultatele observaţiei clinice. Or, ajutaţi de statistica matematică, vom putea cunoaşte, pe baza analizei logice prealabile, gradul acestor erori.

★

Lucrarea de faţă are scopul să indice cele mai simple metode matematice, care însuşite cu uşurinţă de orice medic, să poată fi utilizate în verificarea exactităţii cercetărilor clinice, în medicina experimentală sau în cercetările din domeniul demografiei, morbidităţii şi a dezvoltării fizice a populaţiei.

Din aceste considerente, lucrarea nu se mulţumeşte cu expunerea formulelor matematice, acolo unde este cazul, ci caută să explice cauza pentru care se utilizează aceste formule şi felul cum s-a ajuns la aplicarea acestora în problema cercetată.

Deci scopul acestei lucrări nu este de a perfecţiona cadrele medico-sanitare în probleme de matematică, ci de a le pune la îndemînă elementele de bază absolut necesare pentru verificarea cercetărilor, măsurarea erorilor, calcularea certitudinii diferenţei dintre două valori obţinute, stabilirea existenţei corelaţiilor dintre două sau mai multe fenomene, stabilirea eşantionului necesar, reprezentativ etc.

Însuşind aceste elemente cadrele medico-sanitare vor dispune de încă o metodă — pe lîngă cea experimentală şi de observaţie — în rezolvarea ştiinţifică a cercetărilor efectuate.

2. METODOLOGIA CERCETĂRII

Înainte de a trece la aplicarea metodelor matematice în medicină, vom arăta, în linii generale, metodologia cercetării, etapele mai importante ale acesteia.

Ţinînd seama de cele arătate anterior despre variabilitatea mare a fenomenelor din domeniul biologiei şi îndeosebi a medicinei, cercetătorul nu poate fi indiferent faţă de metodologia utilizată în cercetarea sa şi în prelucrarea materialului obţinut. De utilizarea unei metodologii corecte depind însăşi rezultatele cercetării. Este îndeajuns, de exemplu, ca gruparea materialului, obţinut chiar şi cu o mare precizie, să nu corespundă, pentru ca rezultatele să fie de-a dreptul eronate.

a) Prima fază a oricărei cercetări după *fixarea scopului urmărit* şi punerea la punct a documentaţiei respective o constituie *delimitarea în timp, spaţiu şi volum a obiectului cercetării.*

În fenomenele demografice şi de morbiditate, delimitarea în timp ţine cont de ceea ce este esenţial, şi anume : cunoaşterea tendinţei pe care

o au aceste fenomene, iar pentru cunoaşterea tendinţei, este necesară cel puţin o perioadă de 5 ani sau mai bine de 10 ani.

Cînd efectuarea cercetării an de an întîmpină greutăţi din cauza volumului mare al materialului de examinat, este posibil să se facă cercetările intermitent, la intervale de cîţiva ani. În acest caz este însă nevoie de o perioadă mai lungă de cercetare, şi anume de cel puţin 10 ani.

În cazul dezvoltării fizice, ţinînd seama de volumul mare de cazuri cercetate, intervalele pot fi mai lungi, 5 ani în medie, sau chiar 10 ani.

Delimitarea în spaţiu a obiectului nu întîmpină nici o greutate mai deosebită, depinzînd de scopul urmărit, cunoaşterea situaţiei pe un raion, oraş, pe o circumscripţie etc. Trebuie doar să ţinem seama de faptul că în delimitarea pe teritorii mici probabilitatea fluctuaţiilor este şi mai mare, deci urmărirea tendinţei fenomenelor necesită un interval mult mai lung.

Delimitarea în volum a obiectului este însă mai dificilă, deoarece stabilirea unui volum prea mic poate să ducă la concluzii greşite sau lipsite de importanţă pentru scopul urmărit. Totuşi şi această delimitare în volum se poate face, ţinînd seama de scopul urmărit şi de regulile stabilirii eşantionului reprezentativ.

În cazul cercetărilor clinice şi experimentale, această delimitare este, de asemenea, dificilă.

Metoda de bază în cercetările clinice este observaţia cazurilor individuale, sinteza şi generalizarea rezultatelor obţinute.

Într-o serie de probleme, cum ar fi unele boli cu o mai mare frecvenţă : reumatismul, boala hipertonică etc., delimitarea cercetării în timp, spaţiu şi volum este posibilă, fără riscul de a cerceta un eşantion nereprezentativ.

În alte cazuri, însă, cum ar fi cercetarea într-o serie de boli rare, delimitarea în timp, spaţiu şi volum constituie totdeauna o problemă grea, deoarece concluziile obţinute de multe ori nu sînt reprezentative, iar sporirea numărului de cazuri nu depinde de cercetător.

În aceste condiţii, singura posibilitate pentru a putea ajunge la un eşantion reprezentativ este conlucrarea a mai multor cercetători şi pe o perioadă mult mai mare de timp.

Dar şi în aceste cazuri rare statistica matematică ne pune la dispoziţie o serie de calcule, care ne ajută să stabilim întrucît acest eşantion mic ar putea fi reprezentativ sau nu, deci întrucît el este o parte omogenă a unei colectivităţi.

În medicina experimentală delimitarea cercetării în timp, spaţiu şi volum este de multe ori şi mai îngreunată, cum ar fi, de exemplu, experimentul pe animale de laborator, pentru stabilirea eficienţei unor seruri, vaccinuri sau medicamente.

Numărul de animale de laborator ce servesc la experimentarea eficacităţii sau nocivităţii unor medicamente, ca şi lotul martor de animale constituie un element esenţial în concluziile ce se trag în ceea ce priveşte aceste medicamente.

Concluziile pot fi eronate în cazul în care lotul de animale asupra cărora se efectuează experimentul sau lotul martor nu este suficient de mare.

În unele cazuri, alegerea numărului de animale poate să ducă la rezultate negative în ceea ce privește acțiunea medicamentului sau a serului respectiv, deși în realitate acțiunea lor justă ar fi fost demonstrată în condițiile unei delimitări corecte a numărului de animale de experiență.

În alte cazuri se face risipă, sacrificîndu-se un număr nejustificat de mare de animale, fără ca acest fapt să fi fost necesar.

În toate aceste cazuri, metodele matematice ne pot ajuta, atît în stabilirea eșantionului necesar, cît și în aprecierea rezultatelor obținute.

b) A doua fază a cercetării constă în *efectuarea observației sau a experimentării*.

În clinică, aceasta corespunde observației individuale a bolnavilor și întregistrării acesteia în fișe sau foi de observație, în vederea prelucrării ulterioare. Paralel cu observația, are loc și confruntarea datelor clinice cu cele de laborator, în vederea stabilirii diagnosticului, prognosticului și a comportamentului terapeutic în cazul respectiv.

În medicina experimentală, această fază corespunde însăși experimentelor făcute, cu consemnarea rezultatelor obținute.

În cercetările de teren, în studiul fenomenelor demografice, de morbiditate, această fază corespunde anchetei ce se face pe teren, iar în cazul dezvoltării fizice a populației, măsurătorilor somatometrice ce se fac la copii, cu înregistrarea concomitentă în fișe sau tabele.

Ceea ce este decisiv în această fază, din punct de vedere al înregistrării datelor cercetării, este exactitatea cu care se face observația, experimentul, ancheta, măsurătoarea și înregistrarea acestor date.

Din aceste motive cercetările trebuie făcute de specialiști care cunosc, atît tehnica, cît mai ales importanța și scopul cercetării. Chiar și în acest caz erorile sînt posibile, dar pot fi eliminate la verificarea datelor cercetării.

O problemă esențială în domeniul cercetării este înlocuirea, cît mai mult posibilă, a descrierilor prin date cifrice. Desigur că descrierile nu pot fi înlocuite cu totul, dar în nici un caz ele nu pot înlocui datele cifrice, care fac posibilă o comparare justă și o sinteză a întregii lucrări.

Astfel, în cercetarea clinică, rezultatele în ceea ce privește evoluția bolii : simptomatologia, complicațiile, sfîrșitul bolii prin vindecare sau deces, rezultatele cercetărilor de laborator privind modificarea tabloului sanguin, a urinii, lichidului cefalorahidian etc. nu se pot rezuma numai la descrieri, ci trebuie, pe cît posibil, să li se găsească o exprimare cifrică.

Cu atît mai mult în medicina experimentală sau în cercetările de teren, rezultatele privind eficacitatea, doza utilă din punct de vedere terapeutic, doza letală a medicamentelor, sau înălțimea și greutatea medie la copii, morbiditatea, mortalitatea, decesele la suta de bolnavi trebuie să se sintetizeze în valori cifrice, pe lîngă descrierile care pot și trebuie să însoțească — ca aprecieri — rezultatele cifrice.

În caz contrar, descrierile ar putea compromite cercetarea, dat fiind faptul că ele exprimă părerile subiective ale fiecărui cercetător ; deci înlocuirea cifrelor cu expresii ca : „mai mare", „mai mult", „mai grav", „mai eficace" compromite sinteza și valoarea cercetării, nepermițînd, în nici un caz, o comparabilitate justă a rezultatelor obținute de diverși cercetători.

Cercetarea clinică și experimentală se efectuează, în general, *în mod continuu*, fără perioade de întrerupere. Aceasta este necesar pentru a asigura

condiţiile similare cît mai perfecte, pentru ca rezultatele să nu fie influenţate nici de factorul timp.

Pentru cercetările experimentale îndeosebi este necesar ca durata diferitelor experimentări să nu fie prea mare, căci omogenitatea n-ar mai putea fi asigurată în aceleaşi condiţii.

Există, totuşi, multe situaţii cînd cercetarea pretinde o perioadă lungă şi devine chiar posibilă *cercetarea în etape, periodică*, cu întreruperea, pe un anumit timp, a cercetării. Acesta este cazul în unele cercetări clinice, cum ar fi în bolile rare sau cele care sînt legate de anumite anotimpuri.

În cercetările de teren privind demografia, morbiditatea sau dezvoltarea fizică, forma generală de cercetare (ancheta) se face, de obicei, periodic, nu numai din cauza volumului mare al lucrării, dar şi datorită faptului că starea de sănătate nu se modifică esenţial de la an la an, iar tendinţa acestor fenomene nici nu se poate stabili de la un an la altul.

c) Faza a treia a cercetării o constituie *gruparea şi prelucrarea materialului obţinut prin cercetare*. În general, înregistrarea observaţiei clinice se face pe fişe sau foi de observaţie individuale.

Datele centralizate din foile de observaţie sînt trecute pe tabele care conţin caracteristicile esenţiale pentru problema respectivă : sexul, grupele de vîrstă, starea socială, ocupaţia, formele clinice, complicaţiile, tratamentul, sfîrşitul letal sau vindecarea etc.

Pentru fiecare din caracteristicile de mai sus se defalcă atîtea subgrupe cîte sînt necesare pentru studiul urmărit. Grupele de vîrstă, de exemplu, nu se vor trece mecanic din 5 în 5 ani, ci ţinînd seamă de problema studiată : în poliomielită, grupele de vîrstă la copiii sub 5 ani trebuie defalcate pe ani ; în bolile cardiace, din contră, putem face defalcarea cazurilor pe grupe de vîrstă de cîte 5 ani sau chiar din 10 în 10 ani.

Gruparea materialului de cercetat constituie baza unei juste analize. Tocmai din aceste motive ea trebuie să se facă studiind esenţa materială a fenomenelor cercetate şi, ţinînd seama de aceasta, să se facă aprecierea caracteristicilor esenţiale.

Această problemă nu o poate rezolva decît cercetătorul care cunoaşte problema respectivă, ipoteza de lucru.

Prin gruparea materialului cercetat, acesta se subdivide în părţi care au însuşirea că sînt mai omogene. Dar prin aceasta, colectivitatea de fenomene sau de fiinţe cercetate devine mai puţin numeroasă.

De aci problema ce se pune cercetătorului : pe de o parte, materialul trebuie grupat pe colectivităţi omogene, după caracteristicile esenţiale, iar pe de altă parte, o fărîmiţare prea mare a materialului nu permite tragerea de concluzii juste, numărul mic de cazuri nemaifiind reprezentativ.

Statistica matematică ne ajută să stabilim ce masă de fenomene sau fiinţe trebuie să cuprindă întreaga cercetare şi care este gruparea cea mai mică, dar încă reprezentativă, din colectivitatea cercetată.

Prin gruparea materialului rezultat din cercetarea după caracteristici esenţiale se poate trece la prezentarea lui în tabele, la calcularea de indici şi, în sfîrşit, la analiza calitativă.

Această parte a cercetării, şi anume prelucrarea materialului, cu ajutorul metodelor matematice, formează obiectul lucrării de faţă.

3. LEGEA NUMERELOR MARI

Pentru a lămuri importanţa metodelor matematice şi a înţelege aplicabilitatea ei în cercetarea clinică, experimentală sau de ocrotire a sănătăţii, vom începe cu dezvoltarea noţiunilor generale, urmînd să trecem apoi la aplicarea ei concretă în domeniul cercetărilor medicale.

Pentru a ajunge la definiţia legii numerelor mari să pornim de la cîteva exemple concrete din domeniul medicinei.

EXEMPLUL 1

Se ştie că înălţimea normală la naştere a copiilor este, în medie, de aproximativ 50,5 cm la sexul masculin şi aproximativ 50 cm la sexul feminin.

De unde s-a obţinut această valoare „normală", „medie" a înălţimii la naştere? Din practica zilnică rezultă că înălţimea la naştere (sau greutatea etc.) diferă de la un caz la altul. Această variaţie este mai mare în cazul greutăţii, mai mică în cazul înălţimii, dar ea există.

S-ar putea întîmpla ca din 100 de născuţi vii de sex masculin, nici unul să nu aibă această înălţime „normală", „medie", ci să se abată, în plus sau minus, de la ea.

Cu toată această dispersiune (variaţie, abatere) de la „medie" a cazurilor individuale, există, totuşi, o proprietate comună, şi anume: luînd un număr mai mare de valori ale înălţimii copiilor la naştere, vom constata că ele se situează, în general, în jurul valorii medii. Această valoare „normală" corespunde, deci, valorii medii pe un număr mare de cazuri omogene.

Dacă vom examina înălţimea a 5 nou-născuţi de sex masculin, s-ar putea găsi o medie a înălţimii de 51 cm.

Examinînd înălţimea unui alt lot de 5 nou-născuţi, s-ar putea găsi o medie de înălţime de 49 cm. La un al treilea lot, înălţimea medie ar fi de 52,6 cm ş.a.m.d.

Deci, luînd la întîmplare loturi mici de nou-născuţi, înălţimea va diferi, mai mult sau mai puţin, de la valoarea medie de 50,5 cm. Dar la un număr mare de observaţii: 100, 200 sau 300 vom putea constata că media se apropie, din ce în ce mai mult, de valoarea 50,5 cm. De asemenea, vom constata că o bună parte din cazurile observate se situează în preajma acestei valori.

Să luăm datele privind înălţimea la nou-născuţii dintr-o maternitate din Bucureşti, în anul 1955.

În tabelul I se poate vedea repartiţia celor 258 de nou-născuţi de sex masculin şi tot atîţia de sex feminin, repartiţie făcută fără a fi ales cazurile, ci înregistrînd, la întîmplare, 516 nou-născuţi, repartizaţi în mod egal pe sexe.

Calculînd media aritmetică ponderată a celor două serii de înălţime vom găsi valoarea medie la nou-născuţi de sex masculin de 50,6 cm, iar la nou-născuţi de sex feminin de 49,8 cm.

Din acest tabel se poate constata că faţă de media de 50,6 cm la băieţi şi 49,8 cm la fete, există, în ambele cazuri, o abatere, în plus sau în minus.

Se constată, de asemenea, că, cu toată dispersiunea cazurilor individuale de la medie, există, totuşi, o regularitate şi în această dispersiune, şi anume : atît la băieţi, cît şi la fete, cele mai numeroase cazuri se situează (în exemplul nostru, din punct de vedere al înălţimii) în jurul valorii medii de 50,6 cm, respectiv 49,8 cm.

Tabelul I

Înălţimea la nou-născuţi

Băieţi		Fete	
Înălţime (în cm)	Numărul nou-născuţilor	Înălţime (în cm)	Numărul nou-născuţilor
46	1	42	1
47	5	43	0
48	13	44	2
49	30	45	3
50	89	46	4
51	49	47	12
52	40	48	29
53	20	49	41
54	7	50	75
55	3	51	57
56	0	52	21
57	1	53	8
	258	54	4
		55	1
			258

Astfel, de exemplu, la băieţi, între 49 şi 52 cm inclusiv, se situează 208 cazuri din cele 258, iar la fete, între 48 şi 51 cm inclusiv, se situează 202 cazuri, tot din 258 cîte au fost luate în total.

În mod sugestiv, cele expuse anterior rezultă mult mai bine din graficele denumite : poligonul de frecvenţă şi curba de frecvenţă (graficul din fig. 1 şi 2).

În graficul din fig. 1, denumit poligon de frecvenţă, pe abscisă (OB) se notează intervalele de înălţime, iar pe ordonată (OA) numărul de cazuri corespunzătoare. Rezultă, mult mai clar decît din tabelul I, concentrarea majorităţii cazurilor spre valoarea centrală, media fiind de 50,6 cm, respectiv 49,8 cm, reprezentată în graficul nostru printr-o linie verticală punctată.

În graficul din fig. 2, denumit curbă de frecvenţă, similar cu poligonul de frecvenţă (derivat în mod teoretic din acesta în cazul unor intervale mici pe abscisa graficului, de exemplu din 2 în 2 mm), se poate constata, la fel de clar, distribuţia colectivităţii cercetate.

Luînd la întîmplare, fără să se aleagă, loturi de cîte 5 nou-născuţi, s-au constatat diferenţe mari faţă de media noastră. Astfel, la primul lot de 5 băieţi s-a găsit ca medie de înălţime 48 cm, la al doilea lot 52,5 cm, la al treilea lot 50,5 cm, la al patrulea lot 53 cm etc.

Cu cît numărul de cazuri înregistrate creştea, cu atît media înălţimii se apropia mai mult de 50—51 cm, iar la 194 de cazuri media a ajuns la 50,6 cm. De la acest număr de cazuri în plus, media n-a mai variat

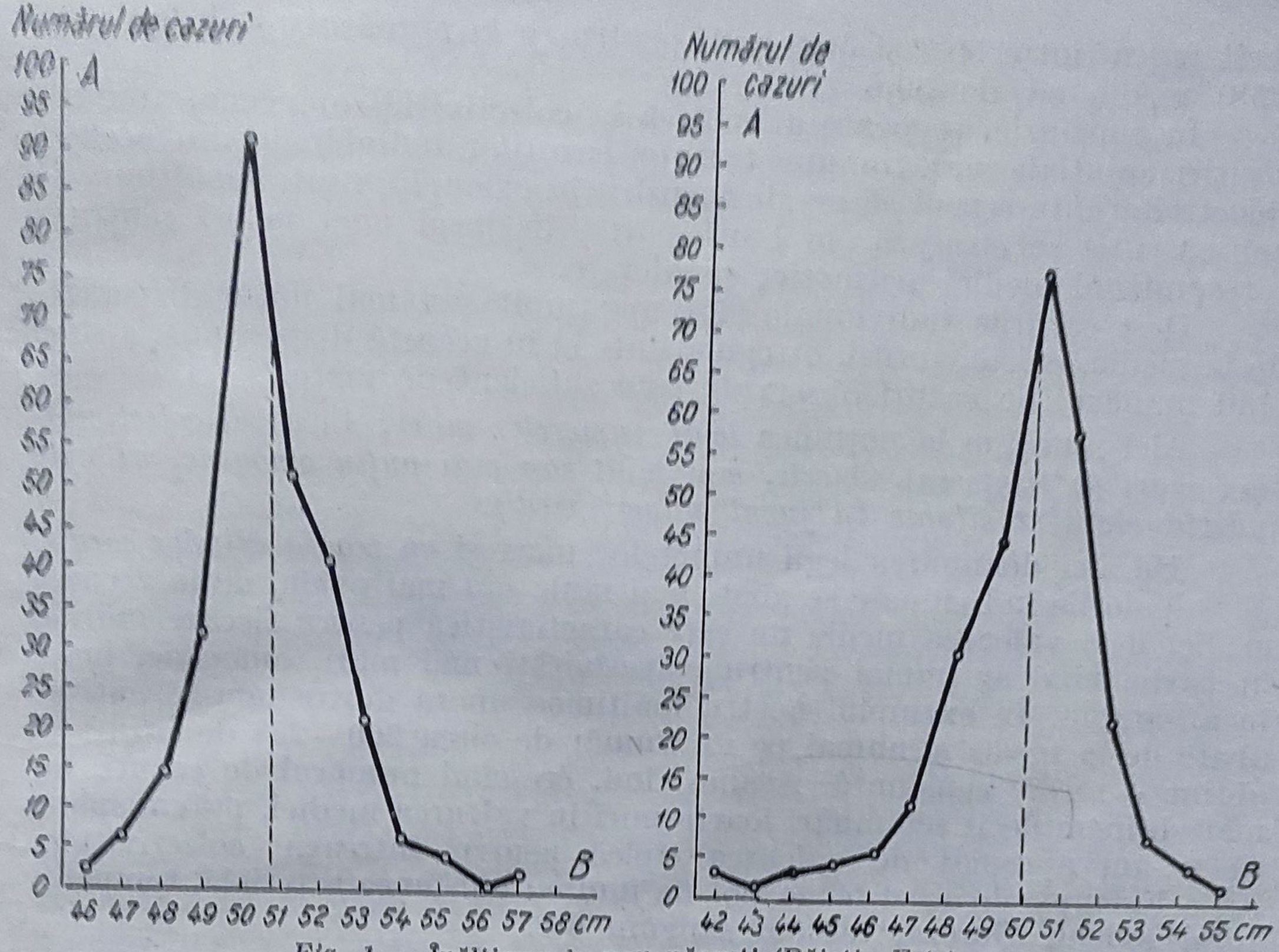

Fig. 1. — Înălțimea la nou născuți (Băieți — Fete).

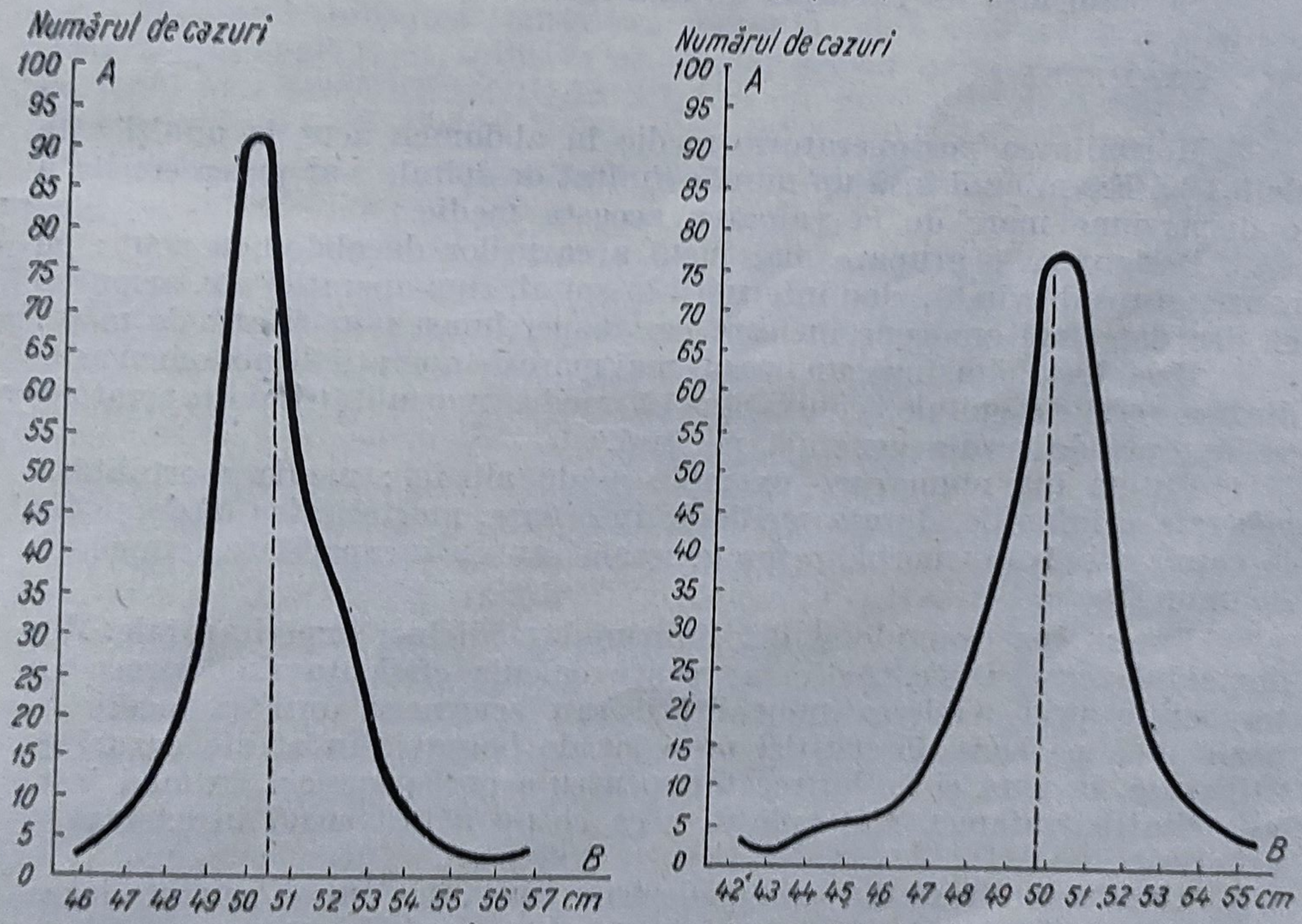

Fig. 2. — Înălțimea la nou născuți (Băieți — Fete).

decît puțin (între 50,3 și 50,8 cm), pentru ca la numărul total de cazuri (258) media să fie 50,6 cm.

În concluzie, se poate afirma că la colectivități omogene, anumite însușiri esențiale sînt comune tuturor cazurilor individuale din această colectivitate, în sensul că aceste însușiri (în exemplul nostru înălțimea la nou-născuți) se situează, în bună parte, în jurul unei valori centrale, corespunzînd mediei aritmetice ponderate.

Deși cazurile individuale sînt mai mult sau mai puțin dispersate de la medie, există, totuși, o regularitate și în această dispersiune, atunci cînd numărul de cazuri observate este suficient de mare.

Aici ajungem la noțiunea *legii numerelor mari: în cazul colectivităților mari de ființe sau obiecte, mai mult sau mai puțin omogene, valorile individuale sînt situate în jurul valorii medii.*

De aici denumirea legii numerelor mari și ca *lege a cifrelor medii.*

Valorile individuale se abat, mai mult sau mai puțin, de la această medie, deci valoarea medie nu este caracteristică pentru fiecare individ în parte, ci doar numai pentru colectivități mai mari, omogene, luate în ansamblu. În exemplul nostru înălțimea unora dintre nou-născuți se abate de la medie și numai pe un număr de circa 200—258 de cazuri se obține o medie constantă. Atunci cînd, crescînd numărul de cazuri, nu mai obținem decît schimbări foarte mici în valoarea mediei, putem spune că ne-am apropiat de valoarea tipică pentru întreaga colectivitate.

Exemple de acest fel se pot da multe : tensiunea arterială, temperatura, pulsul, diureza, tabloul sanguin etc.

Să luăm însă un exemplu de altă natură.

EXEMPLUL 2

Mortalitatea postoperatorie medie în abdomen acut la adulți este de 5,1%. Examinînd însă un număr limitat de spitale s-ar putea constata o dispersiune mare de la valoarea aceasta medie.

Desigur că o grupare mai justă a cazurilor de abdomen acut : pe cauze, grupe de vîrstă, ziua internării în spital, ziua operației etc. ar putea să dea date mai omogene, în care caz dispersiunea n-ar fi atît de mare.

Dacă însumăm, însă, un număr mai mare de operații de abdomen acut dintr-o serie de spitale, constatăm că media mortalității postoperatorii se apropie de media generală pe țară.

Putem lua numeroase exemple și de alt fel : media mortalității generale și infantile, durata medie de internare, mortalitatea la domiciliu la copiii sub 1 an, morbiditatea generală sau cu incapacitate temporară de muncă etc.

Putem lua exemple și din domeniul medicinei experimentale, din domeniul terapiei etc. Astfel se poate calcula eficiența unui procedeu terapeutic prin scăderea mortalității sau scurtarea duratei medii de boală etc., pe baza observației unui lot de bolnavi. În aceste cazuri se utilizează, în plus, și un lot martor, pentru a putea aprecia, în mod just, eficacitatea tratamentului aplicat, ceea ce nu a fost cazul în exemplele anterioare.

Concluziile trase din exemplele anterioare, în ceea ce privește legea numerelor mari, nu trebuie, însă, să ne ducă la fetișizarea acestei legi,

ca fiind o dogmă universal-valabilă. În realitate, esența materială a fenomenelor este aceea care permite aplicarea sau neaplicarea justă a legii numerelor mari.

Astfel faptul că există asemănări fundamentale între copiii nounăscuți de un anumit sex sau grupă de vîrstă sau faptul că există asemănări esențiale între reactivitatea indivizilor de aceeași vîrstă, sex, condiții social-economice etc. face să se poată exprima corect compararea acestor însușiri generale prin legea numerelor mari și nu invers, adică această lege să determine însușirile medii ale unei colectivități.

4. CALCULUL PROBABILITĂȚILOR

La baza legii numerelor mari stă calculul probabilităților. Din exemplul 1 (media înălțimii nou-născuților) rezultă că luînd un singur caz la întîmplare nu putem avea certitudinea că înălțimea acestui copil va fi egală cu media globală. Există, însă, o „probabilitate" ca înălțimea acestuia să se apropie, totuși, de valoarea medie sau să fie chiar identică cu media.

Cu cît numărul de cazuri cercetate va fi însă mai mare, cu atît probabilitatea sau „șansa" de a ne apropia de media globală este mai mare.

Exemplu mult mai sugestiv este aruncarea monedei în aer și probabilitatea („șansa", „speranța") de a cădea pe o anumită față. Un alt exemplu este cel clasic cu urna, în care introducem un număr de bile albe și negre în aceeași proporție, să zicem 500/500, adică într-o proporție de 1/1.

Deși probabilitatea (șansa sau speranța) de a scoate o bilă albă sau neagră este egală (1/2), totuși la un număr scăzut de scoateri vom obține cu totul alte proporții : 4/1 sau 3/1. Cu cît vom mări însă numărul de scoateri (în aceleași condiții, adică reintroducînd în urnă bila scoasă pentru a avea în urnă tot timpul aceeași proporție), cu atît proporția se va apropia de cea normală, adică 1/1.

Această situație simplă se complică însă în cazul în care nu cunoaștem proporția reală a bilelor din urnă, sau, ca în exemplul 1, nu cunoaștem înălțimea pentru toți nou-născuții dintr-o țară sau regiune, raion. În acest caz numai calculul probabilităților, bazat pe examinarea unui număr suficient de mare de cazuri, deci bazat pe un eșantion reprezentativ, va putea elucida (bineînțeles, cu un anumit grad de eroare), valoarea medie a înălțimii.

Trebuie din nou să precizăm, însă, că nici calculul probabilităților care fundamentează legea numerelor mari nu este o dogmă care lămurește și rezolvă totul. După cum s-a mai arătat, analiza esenței materiale a fenomenelor este hotărîtoare în primul rînd și permite ca aceste calcule să fie cît mai corespunzătoare. Din contră, calculul probabilităților nu aduce nimic nou și nu poate rezolva nimic în cazul în care analiza fenomenului respectiv ne arată lipsa de legătură sau lipsa completă de omogenitate dintre elementele colectivității cercetate.

Cu cît calculul probabilităților se va face pe un număr mai mare de cazuri omogene, cu atît și rezultatele vor fi mai sigure, adică mai apropiate de realitate, existînd, totuși, posibilitatea unei erori, eroarea scăzînd o dată cu creșterea numărului de cazuri omogene observate, deci o dată

cu creșterea eșantionului (numărul de cazuri observate din totalul cazurilor existente).

Cu acestea ajungem la definiția probabilității : *probabilitatea de a se produce un fenomen* (în exemplul nostru anterior, probabilitatea de a extrage o bilă albă din urnă) *este raportul dintre numărul de cazuri favorabile producerii fenomenului și totalul cazurilor posibile, cu condiția ca toate posibilitățile să aibă aceeași șansă* (în cazul nostru de a scoate, fie o bilă albă, fie una neagră).

În cazul nostru, probabilitatea (p) de a extrage o bilă albă din urnă este :

$$p = \frac{1}{2} = 0{,}5,$$ adică o șansă din două posibilități.

Exprimat în procente probabilitatea se notează cu P.

$$P = \frac{1}{2} \cdot 100 = 50{,}0\,\%.$$

În cazul aruncării cu zarul (acesta avînd 6 fețe), posibilitatea de a obține una din fețe, de exemplu pe cea de 3, este :

$$p = \frac{1}{6} = 0{,}166,$$ adică o șansă din șase posibilități.

Exprimat în procente probabilitatea aceasta se exprimă în felul următor :

$$P = \frac{1}{6} \cdot 100 = 16{,}6\,\%.$$

Neșansa, adică probabilitatea de a nu se produce evenimentul (contraprobabilitatea) se notează cu q, iar procentul corespunzător prin Q.

În cazul anterior probabilitatea de a nu extrage o bilă albă este :

$$q = \frac{1}{2} = 0{,}5,$$ adică o neșansă din două posibilități.

Procentual această neșansă se exprimă :

$$Q = \frac{1}{2} \cdot 100 = 50{,}0\,\%.$$

În cazul aruncării cu zarul, probabilitatea de a obține dintr-o aruncare una din fețe, pe cea de 3 de exemplu, a fost :

$$p = \frac{1}{6} = 0{,}166,$$ iar $P = 16{,}6\,\%$.

Probabilitatea contrarie, adică neșansa de a obține această valoare este :

$$q = \frac{5}{6} = 0{,}834,$$ iar procentual

$Q = 83{,}4\,\%$.

Din formulele anterioare se poate deduce că suma probabilității și contraprobabilității este egală cu 1.

$p + q = 1,$

adică $\frac{1}{2} + \frac{1}{2} = 1$

sau $0{,}5 + 0{,}5 = 1,$

iar în cazul zarurilor : $\frac{1}{6} + \frac{5}{6} = 1$,

sau $0{,}166 + 0{,}834 = 1$.

Exprimat în procente, vom obține totalul egal cu 100,0%.

$P + Q = 100{,}0\%$, iar în cele două exemple anterioare vom obține :

a) $50{,}0\% + 50{,}0\% = 100{,}0\%$;

b) $16{,}6\% + 83{,}4\% = 100{,}0\%$.

Din formulele de mai sus rezultă, deci, că *suma probabilității și contraprobabilității este egală cu unitatea, iar în procente cu 100,0%.*

Cunoscînd p sau q putem afla valoarea cealaltă :

$p = 1 - q$; $q = 1 - p$.

Aceste formule exprimate procentual vor da :

$P = 100{,}0\% - Q$ sau $Q = 100{,}0\% - P$.

În cazul aruncării cu zarul, cunoscînd probabilitatea $\left(\frac{1}{6}\right)$ găsim contraprobabilitatea : $q = 1 - \frac{1}{6} = \frac{5}{6}$ sau procentual :

$Q = 100{,}0 - 16{,}6 = 83{,}4\%$.

★

În cazul aruncării cu moneda vorbim de *probabilitate simplă,* în sensul că există șansa de a cădea pe o față sau alta, adică de a obține una din cele două posibilități existente.

Probabilitatea compusă se constată în cazul existenței a mai multor probabilități simple, ca, de exemplu, în aruncarea cu zarurile.

Astfel probabilitatea de a obține dintr-o aruncare, sau un 3 sau un 5, este dublă, față de cazurile simple, cînd așteptăm să obținem numai un 3 sau numai un 5.

În calculul probabilităților compuse, în exemplul aruncării cu zarul *șansa de a obține două valori,* fie un 3, fie un 5, *se exprimă prin suma probabilității simple a fiecăreia din valori :*

p_3 sau $p_5 = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$.

Exprimat în procente vom obține :

P_3 sau $P_5 = 16{,}6\% + 16{,}6\% = 33{,}2\%$.

La fel și în cazul cînd așteptăm, să obținem una din trei valori (3 sau 5 sau 6) rezultatul va corespunde sumei probabilității simple a celor trei valori :

p_3 sau p_5 sau $p_6 = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$.

Exprimat în procente obținem :

P_3 sau P_5 sau $P_6 = 16{,}6\% + 16{,}6\% + 16{,}6\% = 50{,}0\%$*.

---

* Prin adunare ar fi 49,8% din cauză că s-a luat 16,6% în loc de 16,66%.

Alta este situaţia cînd aşteptăm să obţinem acelaşi eveniment de mai multe ori. În acest caz *probabilitatea de a se obţine acest eveniment de mai multe ori este mult mai mică, exprimîndu-se nu prin suma probabilităţilor simple, ci prin produsul lor.*

În cazul extragerii bilei albe din urnă probabilitatea de a extrage de două ori la rînd aceeaşi bilă va fi :

$$p \text{ de } 2 \text{ ori} = p \cdot p = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4},$$

adică $0{,}5 \cdot 0{,}5 = 0{,}25$

Exprimat în procente obţinem :

$$P \text{ de } 2 \text{ ori} = \frac{1}{2} \cdot \frac{1}{2} \cdot 100 = \frac{1}{4} \cdot 100 = 25{,}0\,\%.$$

De asemenea, probabilitatea de a obţine de trei ori la rînd acelaşi eveniment, adică de a extrage din urnă de trei ori la rînd numai bile albe, se exprimă tot prin produsul celor trei probabilităţi simple :

$$p \text{ de } 3 \text{ ori} = p \cdot p \cdot p = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8},$$

adică $0{,}5 \cdot 0{,}5 \cdot 0{,}5 = 0{,}125$

Exprimat în procente obţinem :

$$P = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot 100 = \frac{1}{8} \cdot 100 = 12{,}5\,\%.$$

Acestea sînt cele două aspecte mai importante ale calculului probabilităţilor.

Calculul probabilităţilor se aplică, cum se va arăta, în continuare, în numeroase domenii ale medicinei clinice, medicinei experimentale şi ale cercetărilor de teren. El stă la baza calculului dispersiunii de la valorile medii, calculului erorii medii şi a eşantionului reprezentativ etc.

Ţinînd seamă de faptul că fenomenele din domeniul medicinei clinice şi experimentale sînt supuse variabilităţii, calculul probabilităţilor devine o metodă absolut necesară în stabilirea valorilor medii şi a erorii acestor medii.

Astfel stabilirea mortalităţii în anumite boli, frecvenţa posibilă a complicaţiilor, eficienţa probabilă a unor metode terapeutice, a unor seruri, vaccinuri şi medicamente, eficienţă stabilită experimental, necesită neapărat fundamentarea prin calculul probabilităţilor.

Problemele : cît de mare trebuie să fie lotul experimental şi lotul martor în cercetările medicinei experimentale, cîte cazuri clinice trebuie observate pentru a putea trage concluzii valabile pentru descrierea simptomatologiei şi evoluţiei diferitelor boli necesită, de asemenea, aplicarea calculului probabilităţilor.

De asemenea, problemele : care este proporţia reală a morbidităţii generale pe ţară sau în mediul rural, studiat pe baza unui eşantion şi ce mărime trebuie să aibă acest eşantion pentru a fi reprezentativ, care este

stadiul dezvoltării fizice a populației pe baza cercetării de selecție, necesită neapărat aplicarea calculului probabilităților.

În continuare, se vor arăta metodele de studiu privind valorile medii, dispersiunea de la medie, certitudinea și eroarea acestor medii etc.

5. VALOAREA MEDIE

Rezultatele cercetărilor se exprimă, atît cifric, cît și descriptiv. În vederea comparabilității rezultatelor, exprimarea cifrică este absolut necesară. Acest fapt este posibil, deoarece chiar și însușirile calitative pot fi exprimate cifric. În cazul cercetărilor clinice, de exemplu, pot fi exprimate cifric : tensiunea, ritmul cardiac, tabloul sanguin, aciditatea gastrică, densitatea urinii etc.

Există, desigur, unele aspecte ce nu pot fi redate în esența lor prin valori cifrice, cum ar fi, de exemplu, diferitele zgomote cardiace, ralurile bronșice sau intensitatea durerilor gastrice în boala ulceroasă etc.

Aceasta se datorește parțial și aparaturii medicale încă imperfecte. Desigur că în viitor se va putea determina mărimea unor serii întregi de aspecte calitative care astăzi ne sînt descrise prin termeni vagi ca : zgomote surde cardiace, puls filiform etc.

În celelalte aspecte calitative, îndeosebi în simptomatologia subiectivă : durerea, lipsa apetitului, senzația de vomă etc., chiar în lipsa posibilității de a le doza sau măsura, se găsesc posibilități suficiente de exprimare cifrică, bineînțeles sub formă mai puțin exactă. Astfel se poate exprima durerea în felul următor : numărul cazurilor care au durere continuă în regiunea epigastrică, numărul de bolnavi cu dureri intermitente la cîteva ore după mîncare, numărul de bolnavi care au dureri epigastrice doar la palpare etc.

Se pot face, de asemenea, grupări de fenomene cum ar fi durerile epigastrice însoțite sau nu de balonare.

În felul acesta se poate ajunge la grupe de indivizi, fenomene sau însușiri calitative exprimate cantitativ, numeric.

Toate aceste cifre brute sau absolute, ce se obțin în urma cercetării, pot exprima relațiile care există între diferitele fenomene sau colectivități, mai ales în cazul cifrelor mici.

În cazul colectivităților mari comparabilitatea fenomenelor devine aproape imposibilă fără calcularea cifrelor relative.

Cînd afirmăm că într-o secție de chirurgie mortalitatea postoperatorie după apendicectomie este mare, această afirmație o facem ținînd cont de un indice mediu.

Cînd spunem că un bolnav are tensiunea ridicată sau scăzută, această afirmație o facem comparînd-o cu tensiunea arterială „normală", tensiune normală care se exprimă ca media tensiunii arteriale la un număr mare de indivizi sănătoși, de un anumit sex și grupă de vîrstă.

Deci cercetarea trebuie să ducă la concluzii care să arate intensitatea fenomenelor studiate, comparativ cu valorile medii cunoscute anterior sau stabilite chiar în decursul cercetării.

Cînd stabilirea valorii medii nu este posibilă sau eventual nu este necesară, ne putem rezuma la compararea fenomenelor între ele. În acest

ultim caz, însă, nu vom putea aprecia rezultatele cercetării noastre ca avînd valori ridicate sau scăzute, neavînd un etalon, care presupune și o medie cunoscută, „normală".

Mărimile medii pot fi cifre absolute sau cifre relative. Cînd spunem că tensiunea arterială medie la un număr mai mare de indivizi sănătoși, de vîrstă între 30 și 39 de ani, este de 130 mmHg, în acest caz valoarea medie stabilită este o cifră absolută. De asemenea, cînd indicăm că numărul mediu de eritrocite este de 5 000 000 și în acest caz media stabilită este o cifră absolută.

În cazul în care stabilim că mortalitatea postoperatorie prin apendicectomie, la un număr mare de operații, este de 0,03%, în acest caz valoarea medie este o cifră relativă; la fel și în stabilirea valorii globulare, a mortalității prin pneumonie, a mortalității generale sau infantile etc.

Printre valorile medii cele mai mult utilizate sînt:

Media aritmetică simplă și ponderată, care constituie valorile cele mai tipice în aprecierea tendinței centrale a colectivității.

Cu privire la tensiunea arterială la indivizi sănătoși, adulți, media va fi, în general, de 120—130 mmHg. Majoritatea cazurilor individuale se vor situa în jurul acestei valori centrale, medii.

La stabilirea numărului de eritrocite, la persoane sănătoase, media va fi de aproximativ 5 000 000 pe mm^3. Cele mai multe cazuri individuale vor oscila în jurul acestei valori medii, deci vor avea o tendință centrală.

În cazul duratei medii de spitalizare la bolnavii de pneumonie, valoarea medie pe țară este de 10,0 zile. Dacă vom cerceta foile de observație la un număr mai mare de bolnavi de pneumonie, vom constata că cele mai multe cazuri vor avea o durată de 9—11 zile de spitalizare, deci în jurul valorii medii.

În cercetarea bilanțului hidric la muncitorii care lucrează la temperaturi înalte, la stabilirea dozei letale prin diferite substanțe toxice la șoareci, la stabilirea duratei de sîngerare și coagulare, la stabilirea înălțimii și greutății la nou-născut etc. vom obține, pe un număr mai mare de cazuri, valori individuale, oscilînd, de cele mai multe ori, în jurul mediei, avînd deci, în general, o tendință centrală.

Media aritmetică simplă se calculează adunînd toate valorile individuale și împărțind apoi la numărul cazurilor, deci concret *suma valorilor divizată la numărul cazurilor.*

Formula mediei aritmetice simple exprimată prin simboluri va fi:

$$Ma = \frac{\Sigma X}{N}$$

În această formulă a mediei aritmetice simple (Ma) prin X se notează valorile individuale (înălțimea sau greutatea nou-născutului, tensiunea arterială a individului), prin Σ se notează suma. Deci ΣX înseamnă suma tuturor valorilor individuale. Prin N se notează numărul de cazuri observate.

De exemplu, dacă tensiunea arterială a unui grup de 5 persoane ar fi (în mmHg) de : 100, 110, 120, 120, 150, media aritmetică, calculată după formula anterioară, va fi :

$$Ma = \frac{100 + 110 + 120 + 120 + 150}{5} = 120 \text{ mmHg}$$

Cînd cercetăm colectivități mai mari de indivizi, grupîndu-le în clase omogene în vederea unui studiu mai amănunțit, calcularea valorii medii după formula mediei artimetice simple devine inexactă, dat fiind numărul mai mare sau mai mic de cazuri ce corespund valorilor care trebuie adunate. În acest caz se utilizează media aritmetică ponderată, care are următoarea formulă matematică :

$$Mp = \frac{\Sigma X \cdot n}{N}$$

Deci media aritmetică ponderată (Mp) se calculează făcînd suma (Σ) valorilor (X), înmulțite cu frecvența fiecărei valori (n) și împărțite la numărul total de cazuri (N).

Pentru a înțelege mai bine calculul mediei aritmetice ponderate să luăm cîteva exemple :

Dacă un grup de 10 oameni are o tensiune arterială medie de 130 mmHg, iar un alt grup de 20 de oameni are o tensiune arterială de 120 mmHg, pentru calcularea tensiunii arteriale medii pe întreaga colectivitate aplicarea formulei mediei aritmetice simple nu ar fi justă, căci am aduna două grupuri de frecvență diferită (10 cazuri și 20 de cazuri), care au calități diferite (în cazul nostru tensiunea arterială diferită). Am obține, în mod greșit, o tensiune arterială medie de 125 mmHg. Aplicînd însă formula mediei aritmetice ponderate vom obține valoarea medie justă :

$$Mp = \frac{\Sigma X \cdot n}{N} = \frac{130 \text{ mmHg} \cdot 10 + 120 \text{ mmHg} \cdot 20}{10 + 20} =$$

$$= \frac{1\,300 + 2\,400}{30} = \frac{3\,700}{30} = 123 \text{ mmHg}$$

În concretizarea cifrică a formulei matematice am făcut, deci, suma fiecărei valori (în cazul nostru valoarea tensiunii arteriale) a celor două colectivități cercetate, înmulțită cu numărul de cazuri din fiecare grupă (10 cazuri și 20 de cazuri), deci cu frecvența fiecărei valori (*n*). Suma astfel obținută am divizat-o la numărul total de cazuri (30 de cazuri = *N*).

Formula matematică a mediei artimetice ponderate se exprimă, deci, în felul următor : suma produselor fiecărei valori cu frecvența ei, divizată la numărul total de observații.

Să dăm ca un al doilea exemplu seria de variație obținută prin măsurarea înălțimii la nou-născuți.

Aici șirul mai mare de valori cu frecvențe diferite (serie de variație) ar duce la un calcul și mai greșit prin stabilirea valorii mediei cu ajutorul mediei aritmetice simple.

Să luăm înălțimea la nou-născuții de sex masculin.

Tabelul 11

Înălțimea nou-născuților de sex masculin

Înălțimea în cm (X) Coloana 1	Numărul nou-născuților (n) Coloana a 2-a
46	1
47	5
48	13
49	30
50	89
51	49
52	40
53	20
54	7
55	3
56	0
57	1

258 = N = numărul total de nou-născuți

În acest caz media aritmetică simplă, care nu ar ține seama de frecvența fiecărei grupe din seria de variație, ar da o valoare cu totul inexactă : 51,5 cm (prin adunarea coloanei 1 și împărțirea la numărul de intervale de înălțime).

Media aritmetică ponderată ține însă seama de frecvența fiecărei grupe din seria de variație și ca atare rezultatul va fi exact valoarea medie a tuturor observațiilor :

$$Mp = \frac{\begin{array}{c}46 \text{ cm} \cdot 1 + 47 \text{ cm} \cdot 5 + 48 \text{ cm} \cdot 13 + 49 \text{ cm} \cdot 30 + 50 \text{ cm} \cdot 89 + 51 \text{ cm} \cdot 49 + \\ + 52 \text{ cm} \cdot 40 + 53 \text{ cm} \cdot 20 + 54 \text{ cm} \cdot 7 + 55 \text{ cm} \cdot 3 + 57 \text{ cm} \cdot 1\end{array}}{1 + 5 + 13 + 30 + 89 + 49 + 40 + 20 + 7 + 3 + 1} =$$

$$= \frac{13\,064}{258} = 50{,}64 \text{ cm}.$$

Valoarea reală, medie, a înălțimii nou-născuților de sex masculin va fi deci de 50,64 cm și nu de 51,5 cm, cît ne-a dat calcularea ei prin media aritmetică simplă.

Se constată, din exemplele anterioare, că media aritmetică ponderată se apropie de valoarea care are pondere mai mare, adică îi corespunde un număr mai mare de cazuri (o frecvență mai mare).

În exemplul întîi valoarea medie a tensiunii arteriale (123 mmHg) se apropie mai mult de 120 care are o frecvență de 20 de cazuri, decît de valoarea mult mai mare de 130 mmHg care are însă o frecvență mai mică de 10 cazuri.

În al doilea exemplu (tabelul II) valoarea medie a înălțimii nou-născuților de sex masculin (50,64 cm) se cuprinde între valoarea de 50 și 51 cm înălțime, valori care au o cea mai mare frecvență (89 și 49 de cazuri).

Media armonică (M. arm.) se utilizează în calcularea valorii medii, în cazul mai multor mărimi relative. În acest caz utilizarea mediei aritmetice simple nu redă în mod exact valoarea medie, media armonică fiind mult mai apropiată de aceasta.

Formula de calcul a mediei armonice este:

$$M.\ arm. = \frac{\Sigma X}{\Sigma \frac{1}{x} \cdot X}$$

Σ = suma valorilor
X = valorile
x = mărimile relative
$\frac{1}{x}$ = inversul mărimilor relative

Dacă, de exemplu, mortalitatea generală din 3 circumscripții sanitare este de 8,0, 9,0 și 10,0‰, iar numărul de locuitori 4 000, 4 500 și 7 000 și nu cunoaștem cifrele absolute de decedați, media aritmetică simplă este de 9,0‰, în timp ce media armonică, ponderînd valorile relative cu populația la care se referă, va fi:

$$M.\ arm. = \frac{4\,000 + 4\,500 + 7\,000}{\frac{1}{8} \cdot 4\,000 + \frac{1}{9} \cdot 4\,500 + \frac{1}{10} \cdot 7\,000} = \frac{15\,500}{500 + 500 + 700} = 9{,}12‰$$

În realitate media armonică se utilizează rar, în special în stabilirea prețurilor medii.

Media cronologică (M. cron.) se calculează în cazul unei serii de valori, care prezintă o variație în timp (lunară, anuală).

Formula de calcul este următoarea:

$$M.\ cron. = \frac{\frac{X_1}{2} + X_2 + X_3 + X_4 + \frac{X_5}{2}}{N - 1}$$

X (1, 2, 3, 4, 5) = mărimea valorilor la începutul sau sfîrșitul celor patru perioade.

N = numărul de perioade la care ne referim.

Să luăm exemplul următor: numărul de gravide în evidență la început de an a fost de 100, iar la sfîrșitul celor patru trimestre, 110, 124, 132 și 80. Conform mediei aritmetice ar fi fost 109 gravide (media anuală), în timp ce media cronologică este de 114 gravide, conform formulei:

$$M.\ cron. = \frac{\frac{100}{2} + 110 + 124 + 132 + \frac{80}{2}}{5 - 1} = \frac{456}{4} = 114.$$

Media geometrică (M. geom.) se utilizează mai ales în cazul calculării coeficienților de creștere medie a valorilor unei serii cronologice semilo-

garitmice. Media geometrică exprimă mai just valoarea medie, caracterizînd dinamica valorilor acestor serii.

Media geometrică se calculează după formula următoare :

$$M.\ geom. = \sqrt[n]{X_1 \cdot X_2 \ldots X_n}$$

X (1,2, n) = valorile seriei

n = numărul variantelor seriei de valori.

Deoarece la un număr mai mare de variante, extragerea rădăcinii la puterea n este greu de efectuat, problema se rezolvă cu ajutorul logaritmilor, după care utilizînd antilogaritmii se obține media geometrică a seriei de valori.

$$\log M.\ geom. = \frac{\log X_1 + \log X_2 + \ldots \log X_n}{n}$$

Să calculăm media geometrică a următoarelor valori crescînde ale unei serii cronologice : 4, 6, 7, 9, a cărei medie aritmetică este de 6,50

$$M.\ geom. = \sqrt[4]{4.6.7.9} \log\ M.\,geom. = \frac{\log 4 + \log 6 + \log 7 + \log 9}{4} =$$

$$= \frac{0{,}60206 + 0{,}77815 + 0{,}84510 + 0{,}95424}{4} = 0{,}79488$$

M. geom. = Antilog 0,794 88 = 6,22.

Mediana este valoarea care împarte seria de variație în așa fel, încît jumătate din numărul de cazuri se găsește de o parte și de alta a acesteia. Ea se calculează luînd valoarea centrală a seriei de variație, pe baza formulei :

$$Me = \frac{N+1}{2}.$$

Simbolul N corespunde cu numărul de intervale ale seriei de valori, deci nu trebuie confundat cu numărul total de cazuri, care în formulele anterioare s-au notat tot cu N.

Mediana se calculează, deci, luînd valoarea de la mijlocul seriei. În cazul tensiunii arteriale, la un număr de 5 persoane (100, 110, 120, 120, 150 mmHg), mediana va corespunde exact mediei aritmetice, fiind a treia valoare din serie, deci 120 mmHg.

Modulul sau dominanta se calculează și mai simplu, alegînd valoarea căreia îi corespunde frecvența maximă. În cazul anterior al înălțimii nou-născuților de sex masculin, modulul va corespunde deci valorii 50,0 cm, căreia îi corespunde frecvența 89.

Se poate, deci, constata că mediana și modulul nu pot fi luate în considerare, mai ales atunci cînd variabilitatea fenomenelor cercetate este accentuată. Ele totuși au avantajul de a putea fi calculate rapid și simplu pentru a ne putea da seama de valoarea medie. În cazul intervalelor mici de variație, mediana exprimă mai real tendința centrală.

6. METODE SIMPLIFICATE PENTRU CALCULAREA VALORII MEDII

Calculul anterior al mediei aritmetice ponderate în cazul seriilor mari de variație și în cazul utilizării zecimalelor devine destul de migălos, ducînd la posibilitatea de a se face erori. Din aceste motive se utilizează diferite metode de simplificare a calculelor pentru găsirea valorii medii.

Metoda deviației simple se bazează pe proprietatea mediei aritmetice că deviațiile pozitive și negative de la această medie se anihilează, suma lor fiind zero.

Să pornim de la un exemplu simplu. Durata de spitalizare pentru 6 bolnavi de pneumonie a fost, pentru fiecare în parte : 7, 8, 9, 9, 10 și 11 zile.

Valoarea medie a duratei de spitalizare pentru cei 6 bolnavi este de :

$$\frac{7 + 8 + 9 + 9 + 10 + 11}{6} = \frac{54}{6} = 9 \text{ zile.}$$

Deviația cazurilor individuale de la această valoare medie (9 zile) va fi :

$$\begin{aligned} 7 - 9 &= -2 \\ 8 - 9 &= -1 \\ 9 - 9 &= 0 \\ 9 - 9 &= 0 \\ 10 - 9 &= +1 \\ 11 - 9 &= +2 \\ &\overline{0} \end{aligned}$$

Adunînd deviațiile de la medie, cele cu minus și cu plus, vom constata că ele se egalează ($-3 + 3 = 0$)

Pe această proprietate a mediei aritmetice se bazează calculul simplificat al valorii medii.

Cînd nu se cunoaște, în prealabil, că valoarea medie a duratei de spitalizare este de 9 zile, ci se presupune că ar fi 10 zile, urmînd să găsim prin calcul valoarea medie adevărată, procedăm tot prin metoda deviației simple de la medie.

Luînd media ipotetică egală cu 10, deviația cazurilor individuale de la această valoare va fi :

$$\begin{aligned} 7 - 10 &= -3 \\ 8 - 10 &= -2 \\ 9 - 10 &= -1 \\ 9 - 10 &= -1 \\ 10 - 10 &= 0 \\ 11 - 10 &= +1 \\ &\overline{-6} \end{aligned}$$

Adunînd deviațiile de la media ipotetică vom găsi valoarea următoare ($-7 + 1 = -6$).

Din primul moment se poate constata că durata medie de spitalizare nu poate fi egală cu 10 zile, deoarece suma deviațiilor de la această medie ipotetică nu mai dă zero, ci o valoare negativă.

În acest caz putem şti că media reală este sub valoarea acestei medii ipotetice, după cum în cazul unei valori pozitive, media reală ar fi mai mare, peste valoarea mediei ipotetice.

Cum se ajunge de la media ipotetică la media reală? Suma deviaţiilor se împarte la numărul de cazuri şi apoi se adună sau se scade din media ipotetică.

În cazul anterior vom obţine : $-\ 6 : 6 = -\ 1$;

$10 - 1 = 9$ zile.

Să luăm în acelaşi exemplu media ipotetică egală cu 8 zile. Deviaţiile de la această medie vor fi :

$$\begin{aligned} 7 - 8 &= -\ 1 \\ 8 - 8 &= \quad 0 \\ 9 - 8 &= +\ 1 \\ 9 - 8 &= +\ 1 \\ 10 - 8 &= +\ 2 \\ 11 - 8 &= +\ 3 \\ \hline &+\ 6 \end{aligned}$$

Suma deviaţiilor ($+ 6$), împărţită la numărul cazurilor spitalizate (6), va da valoarea $+ 1$, care adunată cu media ipotetică (8) va da media adevărată a duratei de spitalizare (9 zile).

Care este simplificarea în metoda aceasta a deviaţiei simple? Aceea că în locul calculelor greoaie în cazul valorilor mari, acestea se simplifică prin utilizarea cifrelor mult mai mici.

Metoda momentelor. Calculul momentului 1. În cazul calculării mediei aritmetice ponderate, cînd există valori mari sau cifre zecimale care trebuie înmulţite cu frecvenţele lor şi apoi împărţite la numărul total de observaţii, operaţia devine deosebit de dificilă. În acest caz recurgem la o metodă simplificată a calculării mediei aritmetice ponderate, şi anume metoda momentelor.

Luăm ca exemplu seria de variaţie a înălţimii nou-născuţilor de sex masculin, expusă şi în tabelul I şi II. Prima coloană cuprinde seria de variaţie a valorilor de înălţime (X), iar a doua coloană frecvenţele (numărul de cazuri, notat cu n) ce corespunde fiecăreia din valori (tabelul III). Coloana a 3-a corespunde produsului valorilor cu frecvenţele respective, adică produsul datelor corespunzătoare din coloana I şi a II-a (X.n).

Dacă adunăm coloana a 3-a şi împărţim la numărul total de cazuri (N) vom obţine media aritmetică ponderată $\left(Mp = \frac{\Sigma X \cdot n}{N} \right)$, pe care am calculat-o deja anterior, egală cu 50,64 cm. Pentru a înlocui acest calcul, care poate deveni destul de greoi în cazul valorilor mari sau cu zecimale, precum şi în cazul frecvenţelor mari, recurgem la calculul momentului 1.

Acesta se efectuează în felul următor : alegem, în mod arbitrar, ca medie ipotetică, una dintre valorile înălţimii, şi anume aceea căreia îi corespunde frecvenţa cea mai mare (modulul) ; în cazul nostru vom alege ca medie ipotetică înălţimea de 50 cm. În dreptul acestei valori notăm, în coloana a 4-a, zero corespunzînd originei şi apoi completăm coloana, în sus, spre valorile mai mici de înălţime, notînd cu minus : -1, -2, -3, -4 etc., iar în jos notînd cu plus : $+1$, $+2$, $+3$, $+4$ etc. în progresie

aritmetică. Această coloană a deviației de la medie, sau mai bine-zis de la origine o notăm cu X′. În coloana următoare (a 5-a) trecem produsul dintre coloana a 2-a și a 4-a, adică produsul deviației de la medie cu frecvențele corespunzătoare (X′.n).

Vom obține în coloana a 5-a, rezultatele cu semnul algebric minus (—), iar în partea de jos sub origine, spre valorile mari ale seriei de variație, rezultate cu semnul algebric plus (+).

Adunăm separat cele două grupe de produse și apoi scădem suma valorilor cu minus din suma valorilor cu plus, rezultatul împărțindu-l la totalul cazurilor (N).

Formula momentului 1 va fi deci următoarea :

$$M_1 = \frac{\Sigma X' \cdot n}{N}$$

Înlocuind simbolurile formulei momentului 1 cu datele rezultate din tabelul III vom obține :

$$M_1 = \frac{-75 + 239}{258} = \frac{164}{258} = 0{,}64 \text{ cm}$$

Adunăm această valoare a momentului 1 cu media ipotetică aleasă în mod arbitrar și vom obține media aritmetică ponderată.

$$Mp = 50{,}0 \text{ cm} + 0{,}64 \text{ cm} = 50{,}64 \text{ cm}.$$

Tabelul III

Înălțimea nou-născuților de sex masculin

Valorile înălțimii în cm (X)	Frecvențele (numărul cazurilor) (n)	Produsul valorilor cu frecvențele (X.n)	Deviația de la medie (X′)	Produsul deviației de la medie cu frecvențele (X′.n)	
1	2	3	4	5	
46	1	46	—4	— 4	
47	5	235	—3	—15	
48	13	624	—2	—26	
49	30	1 470	—1	—30	—75
50	89	4 450	0	0	
51	49	2 499	1	49	
52	40	2 080	2	80	
53	20	1 060	3	60	
54	7	378	4	28	
55	3	165	5	15	
56	0	0	6	0	
57	1	57	7	7	+239
Total :	258 = N	13 064 = ΣX . n	—	164 = ΣX′ . n	

Din tabelul III se poate constata felul în care utilizarea momentului 1 uşurează calcularea mediei. Astfel în loc de a înmulţi coloana 1 (X) cu coloana a 2-a (n) obţinînd astfel coloana a 3-a ($X.n$) cu cifre mari, care însumate trebuie împărţite la numărul total de cazuri (N), vom înmulţi doar coloana a 4-a (X') cu coloana a 2-a (n). Deci înmulţirea se va face cu 4,3,2,1 etc. în loc de 46, 47, 48, 49 etc. împărţirea la numărul total de cazuri (N) este, de asemenea, uşurată prin metoda momentului 1, căci în loc de a împărţi 13 064 la 258, ca în calculul simplu al mediei aritmetice ponderate, în cazul momentului 1 simplificăm calculul prin scăderea valorilor cu minus din cele cu plus şi apoi efectuăm împărţirea, adică împărţim 164 la 258, rezultatul de 0,64 cm adunîndu-l la media ipotetică (50,0 cm).

Specificăm că se poate alege ca medie ipotetică oricare dintre valorile de înălţime, rezultatul final fiind identic, dar pentru a uşura cît mai mult calculul se alege, aşa cum s-a arătat, valoarea cît mai apropiată de mediană şi modul.

Astfel dacă în exemplul nostru am lua ca medie ipotetică nu 50 cm, ci 51 cm am ajunge la aceeaşi valoare medie (50,64 cm), utilizînd calculul momentului 1, cu singura diferenţă că în acest caz momentul 1 are o valoare negativă, ce trebuie scăzută din media ipotetică (tabelul IV).

Tabelul IV

Înălţimea nou-născuţilor de sex masculin

Valorile înălţimii în cm (X)	Frecvenţele (numărul cazurilor) (n)	Deviaţia de la medie (X′)	Produsul deviaţiei de la medie cu frecvenţele (X′.n)	
1	2	3	4	
46	1	−5	− 5	
47	5	−4	−20	
48	13	−3	−39	
49	30	−2	−60	
50	89	−1	−89	−213
51	49	0	0	
52	40	1	40	
53	20	2	40	
54	7	3	21	
55	3	4	12	
56	0	5	0	
57	1	6	6	+119
Total :	258 = N	—	−94 = ΣX′ . n	

Din tabelul IV, spre deosebire de tabelul III, am eliminat o coloană, cea care corespunde produsului dintre valorile seriei cu frecvenţele (X.n), nefiindu-ne necesară pentru a demonstra avantajele calculării mediei prin metoda momentului 1.

Formula momentului 1 se calculează ca şi în exemplul anterior :

$$M_1 = \frac{-213 + 119}{258} = \frac{-94}{258} = -0{,}36 \text{ cm.}$$

Media aritmetică ponderată căutată se calculează scăzînd momentul 1 din media ipotetică.

$$51 \text{ cm} - 0{,}36 \text{ cm} = 50{,}64 \text{ cm}.$$

Deci rezultatul este identic în toate împrejurările, indiferent de media ipotetică aleasă.

Pentru uşurarea calculului se ia, însă, ca medie ipotetică, totdeauna, valoarea care are o frecvenţă mai mare, sau care se situează în jurul medianei.

Există cazuri cînd intervalele valorilor seriei (raţia de creştere) nu sînt de 1 cm, ci de 2 sau 3 cm, sau —din contra— de 0,5 cm sau 0,3 cm.

În aceste cazuri, înainte de a trece la calculul mediei aritmetice ponderate trebuie să înmulţim momentul 1 cu 2 sau 3, fie cu 0,5 sau cu 0,3.

Metoda calculării mediei prin îndoirea seriei de variaţie. În cazul frecvenţelor mari operaţiile matematice pot fi simplificate la maximum prin utilizarea în calculul momentelor a metodei îndoirii seriei de variaţie.

Această metodă constă în scăderea frecvenţelor (n) ce corespund deviaţiilor negative de la medie (X′), din frecvenţele ce corespund deviaţiilor pozitive de la medie.

Să luăm exemplul anterior (tabelul III) cu înălţimea nou-născuţilor de sex masculin, desfăşurînd, de data aceasta, orizontal coloana a 2-a, ce corespunde frecvenţei (n) şi coloana a 3-a, ce corespunde deviaţiilor de la medie (X′):

−4	−3	−2	−1	0	1	2	3	4	5	6	7
1	5	13	30	89	49	40	20	7	3	0	1

Se ştie că înălţimea de 50 cm a fost considerată ca medie ipotetică (coloana 1 din tabelul III), acestei medii corespunzîndu-i o frecvenţă de 89 de cazuri şi deviaţia zero.

Îndoim seria frecvenţelor în jurul unui ax fix, care trece prin media ipotetică, adică în dreptul frecvenţei de 89 şi deviaţiei zero şi facem scăderile frecvenţelor, notate cu minus şi cu plus, după cum ele corespund deviaţiilor negative sau pozitive. Frecvenţele aflate în dreptul deviaţiilor de la medie negative vor împrumuta şi ele semnul algebric minus (—), iar frecvenţele aflate în dreptul deviaţiilor de la medie pozitive vor avea semnul algebric plus (+).

După îndoirea seriei de frecvenţe scădem frecvenţele negative din cele pozitive:

89, 49, 40, 20, 7, 3, 0, 1

0 −30 −13 −5 −1

89 19 27 15 6 3 0 1

După ce am scăzut frecvenţele, trecem la calculul momentului 1, făcînd înmulţirea cu deviaţiile de la medie.

0 1 2 3 4 5 6 7

89 19 27 15 6 3 0 1

Rezultă din înmulțire :

0, 19, 54, 45, 24, 15, 0, 7, care însumate dau 164

Momentul 1 va fi : $\frac{164}{258} = 0{,}64$.

Adunînd 0,64 cm cu media ipotetică vom găsi media aritmetică ponderată căutată : 50,0 cm + 0,64 cm = 50,64 cm.

La o privire mai atentă se poate descoperi imediat asemănarea dintre calcularea mediei cu utilizarea îndoirii seriei de variație și calculul efectuat fără îndoirea seriei.

În calculul simplu se înmulțește deviația de la medie cu frecvențele corespunzătoare și apoi se face scăderea produselor negative din cele pozitive. În calculul îndoirii seriei de variație, se scad mai întîi frecvențele ce corespund deviațiilor negative din frecvențele ce corespund deviațiilor pozitive și de-abia după această simplificare radicală se efectuează înmulțirea cu deviațiile de la medie.

Deci în cadrul frecvențelor oricît de mari, metoda îndoirii seriei de variație permite calcularea din memorie a majorității înmulțirilor ce trebuie efectuate.

7. DISPERSIUNEA DE LA MEDIE

Valorile medii caracterizează însușirile generale, comune întregii colectivități. Ele însă ascund, în același timp, o serie de însușiri individuale, uneori de importanță deosebită.

Astfel durata medie de spitalizare în pneumonie ascunde durata mai scurtă sau mai îndelungată a cazurilor individuale, deși o prelungire a spitalizării unor bolnavi poate să denote complicații ivite, lipsa unui tratament adecvat etc.

În febra tifoidă, de exemplu, o durată medie mare de depistare (nr. zile de la îmbolnăvire la descoperire) de 7 zile denotă, în general, o lipsă a rețelei sanitare, precum și un nivel necorespunzător de cultură sanitară a populației. Dar analiza cazurilor individuale de febră tifoidă ne poate arăta că unii dintre bolnavi au fost depistați în ziua a 5-a, iar alții în ziua a 9-a și a 10-a, ceea ce constituie o lipsă și mai serioasă, mai ales în cazul cînd ei provin din focare deja cunoscute.

Mortalitatea postoperatorie medie în apendicectomie de 0,03% poate fi considerată ca un indice bun, deoarece aceasta înseamnă doar 3 decese la 10 000 operații, dar studiul acestor cazuri de deces prin apendicectomie poate să dezvăluie unele lipsuri, constînd în depistarea tardivă, în tehnica operatorie sau grija postoperatorie deficientă.

Mortalitatea infantilă poate fi scăzută pe țară, dar o serie de raioane pot prezenta, totuși, o dispersiune în plus sau minus destul de mare de la această valoare medie.

De asemenea, în cazul înălțimii sau greutății medii la naștere, față de valoarea medie, există variații mari de la un caz la altul, o dispersiune importantă față de medie.

Se pot enumera multe exemple de acest fel și în domeniile cercetării clinice sau experimentale, ceea ce demonstrează că valoarea medie este

insuficientă pentru caracterizarea colectivităților, pentru cunoașterea fenomenelor. Cifrele medii maschează, de cele mai multe ori, realitatea. Ele redau doar un singur caracter, comun întregii colectivități. O serie de alte caracteristici esențiale rămîn necunoscute prin utilizarea valorilor medii.

Din aceste motive se utilizează măsurători ale dispersiunii (ale abaterii, variației) de la valoarea medie.

Există numeroase măsurători ale dispersiunii de la medie. Vom enumera doar unele dintre acestea.

Amplitudinea este măsurătoarea extremelor unei dispersiuni, corespunzîndu-i intervalul dintre valoarea maximă și minimă a unei serii de variație.

În exemplul anterior înălțimea la nou-născuți de sex masculin, amplitudinea este de 11 cm, adică diferența înălțimii maxime și minime (57 cm și 46 cm).

Amplitudinea ne arată caracterul dispersiunii, limitele ei maxime și minime. Avînd două colectivități de copii de același sex și grupă de vîrstă, una cu amplitudinea de 11 cm și cealaltă de 14 cm, vom putea conclude că prima colectivitatea este mai omogenă și că media ei caracterizează mai bine întreaga colectivitate.

Deviația cvartilei. Cvartilele precizează mai just dispersiile variantelor față de valoarea lor medie.

Ele împart seria variantelor în patru părți egale.

Cvartila inferioară delimitează primul sfert al seriei de celelalte trei sferturi. Cvartila mijlocie este egală cu mediana, împărțind seria variantelor în jumătate. Cvartila superioară desparte primele trei sferturi din seria variantelor, de ultimul sfert.

Jumătate dintre valorile variantelor, cuprinse între cvartila superioară și cea inferioară se numește deviația cvartilei de la medie și are următoarea formulă:

$$Q = \frac{Q_3 - Q_1}{2}$$

Să luăm un exemplu privind perimetrul toracic la un număr de 8 persoane și să construim distribuția variantelor (valorile sînt redate în cm):

91, 92, 93, 94, 95, 97, 101, 102.

Valoarea medie, adică media aritmetică a perimetrului toracic, va fi de 95,6 cm.

Cvartila inferioară va fi egală cu: $\frac{92 + 93}{2} = 92{,}5$ cm.

Cvartila superioară va fi egală cu: $\frac{97 + 101}{2} = 99$ cm.

Cvartila mijlocie, adică mediana, va fi egală cu:

$\frac{94 + 95}{2} = 94{,}5$ cm.

Deviația cvartilei va fi egală cu: $\frac{99 - 92{,}5}{2} = \frac{6{,}5}{2} = 3{,}25$ cm.

Deviația de la mediană, deci de la cvartila mijlocie la cvartila inferioară sau superioară, se numește *deviație probabilă* și în situația de mai sus este egală, în primul caz, cu 2 cm, iar în al doilea caz cu 4,5 cm.

Aceasta înseamnă că valorile peste mediană variază mult mai mult decît valorile sub mediană.

Decilele și centilele împart distribuția de valori în 10 sau 100 de părți egale. Se utilizează mai ales în cazul seriilor mari de variante a valorilor.

Amplitudinea și deviația cvartilei nu spun însă nimic asupra caracterului dispersiunii, ele descriind doar limitele ei. Din aceste motive pentru a putea caracteriza în mod just o colectivitate, pentru a vedea dacă ea este omogenă, dacă selecțiunea este bine făcută sau dacă numărul de observații este suficient, se utilizează o altă măsurătoare a dispersiunii, și anume : deviația-tip.

Deviația-tip sau deviația standard (abatere tip) se notează cu litera grecească sigma (σ). Ea se calculează mai simplu din seria de variație cu ajutorul momentelor. Momentul 1 am văzut deja cum se calculează ; rămîne să vedem cum se calculează momentul al 2-lea. Cunoscînd momentul 1 și al 2-lea deviația-tip devine ușor de calculat, pe baza formulei următoare :

$$\sigma = \pm \sqrt{M_2 - (M_1)^2}$$

Deci deviația-tip se calculează extrăgînd rădăcina pătrată din diferența dintre momentul al 2-lea și pătratul momentului 1.

Momentul al 2-lea se calculează făcînd produsul dintre deviațiile de la medie, ridicate la pătrat, cu frecvențele corespunzătoare, după care rezultatul se divide la numărul total de cazuri observate. În vederea calculului să desfășurăm din nou tabelul conținînd seria de variație a înălțimii la nou-născuți de sex masculin (tabelul V).

Primele patru coloane sînt deja cunoscute. În coloana a 5-a se trec deviațiile de la medie (X') din coloana a 3-a, ridicate la pătrat, iar în coloana a 6-a, datele din coloana a 5-a, înmulțite cu frecvențele (n).

Tabelul V

Înălțimea nou-născuților de sex masculin

Valorile înălțimii în cm (X)	Frecvențele (numărul cazurilor) (n)	Deviația de la medie (X')	Produsul deviației de la medie cu frecvențele ($X'.n$)	Deviația medie la pătrat $(X')^2$	Produsul pătratului deviației de la medie cu frecvențele $(X')^2.n$
1	2	3	4	5	6
46	1	−4	− 4	16	16
47	5	−3	−15	9	45
48	13	−2	−26	4	52
49	30	−1	−30	1	30
			−75		
50	89	0	0		0
51	49	1	49	1	49
52	40	2	80	4	160
53	20	3	60	9	180
54	7	4	28	16	112
55	3	5	15	25	75
56	0	6	0	36	0
57	1	7	7	49	49
			+239		
Total :	258 = N	—	164 = $\Sigma X'.n$	—	768 = $\Sigma(X')^2.n$

Momentul 1 a mai fost calculat. Repetăm totuşi calculul pentru a cuprinde la un loc toate operaţiile în vederea aflării deviaţiei-tip.

$$M_1 \text{ (momentul 1)} = \frac{\Sigma X' \cdot n}{N} = \frac{\text{coloana a 4-a}}{\text{coloana a 2-a}} = \frac{-75 + 239}{258} = 0{,}64 \text{ cm}$$

Pe baza momentului 1 calculăm media aritmetică ponderată: 50 cm + 0,64 cm = 50,64 cm.

Trecem la calculul momentului al 2-lea, care se obţine împărţind suma din coloana a 6-a din tabelul V la numărul total de cazuri.

$$M_2 \text{ (momentul al 2-lea)} = \frac{\Sigma \cdot (X')^2 \cdot n}{N} = \frac{\text{coloana a 6-a}}{\text{coloana a 2-a}} = \frac{768}{258} = 2{,}98.$$

Cunoscînd momentele 1 şi al 2-lea putem afla deviaţia-tip (σ) extrăgînd rădăcina pătrată din momentul al 2-lea, din care se scade, în prealabil, momentul 1 la pătrat.

$$\sigma = \pm \sqrt{\frac{\Sigma (X')^2 \cdot n}{N} - \frac{(\Sigma X' \cdot n)^2}{N^2}} = \pm \sqrt{\frac{N \cdot \Sigma (X')^2 \cdot n - (\Sigma X' \cdot n)^2}{N^2}}$$

$$\sigma = \pm \sqrt{M_2 - (M_1)^2} = \pm \sqrt{2{,}98 - 0{,}41} = \pm \sqrt{2{,}57} \doteq \pm 1{,}60 \text{ cm}.$$

Aceasta înseamnă că deviaţia-tip este de $\pm$1,6 cm.

Deviaţia-tip (σ) *este un indicator foarte sensibil, care ne arată gradul de dispersiune de la medie a unei serii de valori. Ea ne arată gradul de omogenitate sau dispersiune a colectivităţii respective, deci dacă eşantionul a fost sau nu bine ales.*

Cu cît deviaţia-tip este mai mare, cu atît media caracterizează mai puţin bine, mai inexact, colectivitatea aleasă.

Dacă aceeaşi colectivitate este cercetată după o anumită perioadă de timp şi cu această ocazie deviaţia-tip calculată este mai mică decît deviaţia-tip calculată cu prima ocazie, putem să presupunem, cu un mare grad de probabilitate, că situaţia s-a îmbunătăţit din punct de vedere al omogenităţii, colectivitatea de copii, de exemplu, este mai armonios dezvoltată, fără diferenţe mari pozitive şi negative, deci fără o dispersiune mare de la valoarea medie, care în acest caz caracterizează bine întreaga însuşire a colectivităţii.

Media caracterizează mai bine valorile care se abat cu un σ în plus sau minus de la medie, deci este tipică pentru aceste valori, mai puţin tipică este pentru valorile ce se abat faţă de medie cu $\pm 2\sigma$ şi caracterizează şi mai puţin valorile ce se abat cu $\pm 3\sigma$.

În toate împrejurările în care studiem colectivităţi sau fenomene care au un caracter de variabilitate, indiferent că este vorba de înălţime la nou-născuţi, mortalitatea infantilă, durata medie de spitalizare, decesele postoperatorii, evoluţia şi decesele în diferite boli, metabolismul bazal în legătură cu efortul, eficacitatea unor medicamente sau vaccinuri etc. se utilizează calculul mediei şi al dispersiunii, fie în felul arătat în tabelul V, fie mult mai simplu, pe baza calculului probabilităţii, în cazul frecvenţelor relative, aşa cum se va arăta mai departe.

Să luăm încă un exemplu privind înălţimea la nou-născuţi de sex masculin, pe un eşantion de 262 de cazuri (tabelul VI).

Completînd tabelul tip şi luînd ca medie ipotetică 50,0 cm, obţinem datele următoare :

Tabelul VI

Înălţimea născuţilor de sex masculin

Înălţimea în cm (X)	Frecvenţa (numărul nou-născuţilor) (n)	Deviaţia de la medie (X')	Deviaţia medie ori frecvenţa (X'·n)	Deviaţia medie la pătrat (X')²	Deviaţia medie la pătrat ori frecvenţa (X')²·n
1	2	3	4	5	6
47	3	−3	− 9	9	27
48	14	−2	− 28	4	56
49	41	−1	− 41	1	41
			− 78		
50	112	0	0	0	0
51	50	1	50	1	50
52	22	2	44	4	88
53	11	3	33	9	99
54	7	4	28	16	112
55	2	5	10	25	50
			+165		
Total :	262 = N	—	87 = ΣX'.n	—	523 = Σ(X')².n

Calculăm momentul 1 şi al 2-lea, trecînd apoi la aflarea mediei aritmetice ponderate şi a deviaţiei-tip.

$$M_1 \text{ (momentul 1)} = \frac{\text{coloana a 4-a}}{\text{coloana a 2-a}} = \frac{-78 + 165}{262} = \frac{87}{262} = 0{,}33 \text{ cm.}$$

Media aritmetică ponderată = 50,0 cm + 0,33 cm = 50,33 cm.

$$M_2 \text{ (momentul al 2-lea)} = \frac{\text{coloana a 6-a}}{\text{coloana a 2-a}} = \frac{523}{262} = 1{,}99.$$

$$\text{Deviaţia-tip } \sigma = \pm\sqrt{1{,}99 - (0{,}33)^2} = \pm\sqrt{1{,}88} = \pm 1{,}37 \text{ cm.}$$

Dacă vom compara valorile rezultate din calculele efectuate la cele două eşantioane, care se referă la aceeaşi problemă (înălţimea nou-născuţilor de sex masculin), vom constata următoarele :

— Valoarea medie a înălţimii, măsurată prin media aritmetică ponderată, este sensibil egală la cele două eşantioane cu un număr aproape egal de cazuri. La primul eşantion media înălţimii este de 50,64 cm, iar la al doilea 50,33 cm.

— Amplitudinea primei colectivităţi este de 11 cm, iar a celei de a doua de 8 cm. De aici se poate constata că, în general, cazurile sînt dispersate pe un interval mai mare la prima colectivitate decît la cea de a doua.

Desigur că de aici încă nu rezultă cît este deviaţia, în medie, a cazurilor, care poate varia indiferent de amplitudine.

— Deviaţia-tip la cele două colectivităţi diferă mult între ele, cu circa 17%, fiind de ±1,60 cm la primul eşantion şi ±1,37 la cel de-al doilea.

Se poate deci deduce că a doua colectivitate este mai omogenă decît prima, deci că valoarea ei medie caracterizează mai just întreaga colectivitate decît în cazul primei colectivităţi.

Putem afirma că a doua colectivitate este mai uniform dezvoltată ca înălțime, cazurile individuale fiind, în medie, mai puțin dispersate ca în prima colectivitate.

Situația ar fi similară dacă cele două colectivități ar fi fost eșantioane obținute la perioade diferite, primul eșantion în anul 1950, de exemplu, iar al doilea eșantion în anul 1960. Am fi putut trage concluzii privind

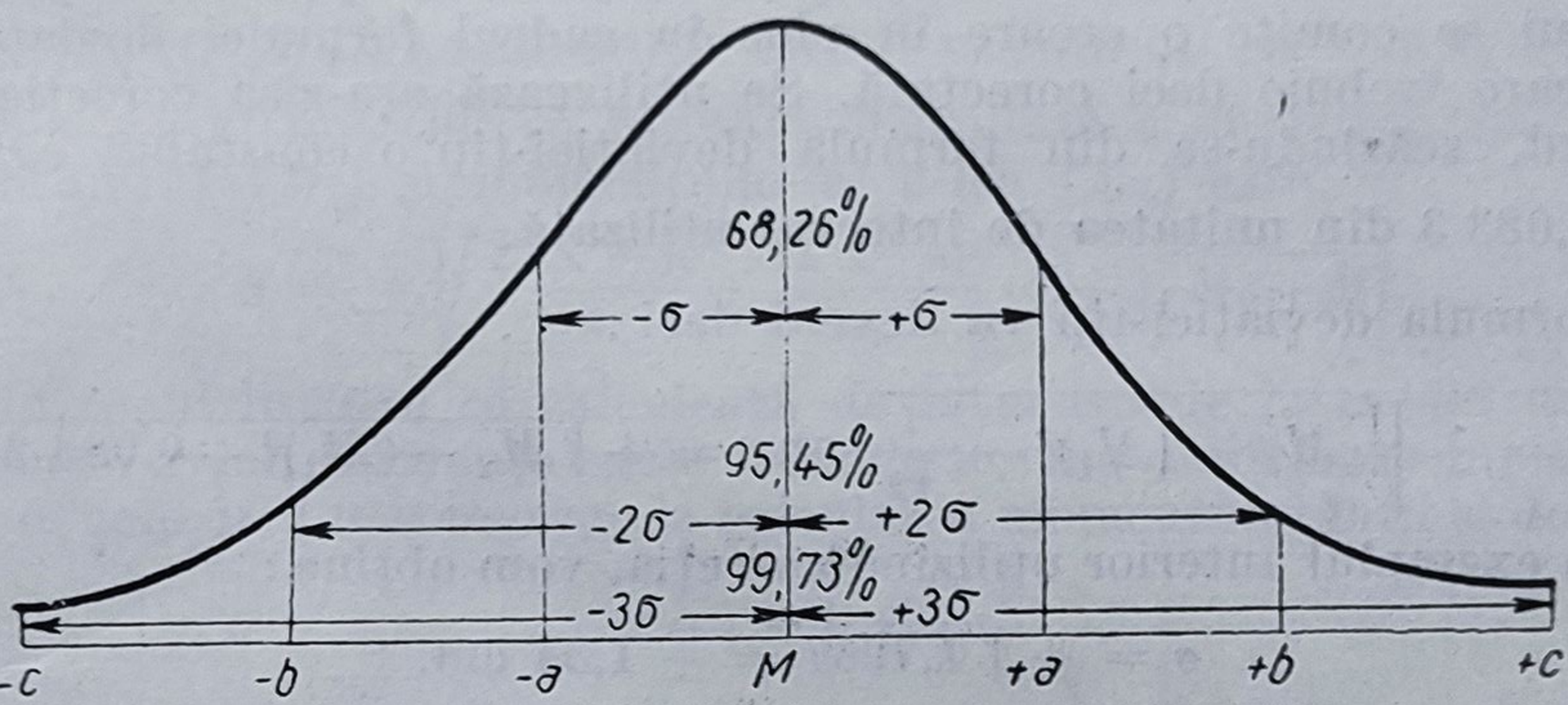

Fig. 3. — Curba Laplace-Gauss.

dinamica dezvoltării fizice la nou-născuți (înălțimea), concluziile fiind aceleași ca și cele specificate anterior.

Deviația-tip este, deci, cu atît mai mare cu cît fenomenul va fi mai variabil, mai dispersat.

Dispersiunea de la medie se poate reprezenta sugestiv prin curba ideală de variație, corespunzînd unei distribuții simetrice și continue (curba Laplace-Gauss).

Curba ideală de variație este simetrică față de ordonata maximă, notată cu linia *M*, corespunzînd, în cazul nostru, mediei aritmetice ponderate.

Pe bază de calcule s-a stabilit, că în cazul curbei ideale sau normale de variație, în partea ariei totale cuprinse între medie, adică ordonata maximă și ordonata a cărei abscisă este egală cu $\pm\sigma$ se cuprinde 68,26% din totalul valorilor unei serii de variație. Între medie și $\pm 2\sigma$ se cuprinde 95,45% din totalul valorilor seriei, iar între medie și $\pm 3\sigma$ se găsesc 99,73% din totalul valorilor seriei de variație.

În calculele noastre nu avem nevoie de o exactitate mai mare în măsurarea dispersiunii decît de $\pm 3\ \sigma$, căreia îi corespunde, într-o distribuție normală, aproape totalitatea cazurilor, cu excepția a 0,27 cazuri dintr-o sută, ceea ce este neglijabil.

Chiar dacă am calcula dispersiunea de la medie cu $\pm 4\sigma$ n-am obține o exactitate cu mult mai mare ($+4\sigma = \pm$ 99,9936%, iar $\pm 5\sigma =$ $= \pm$ 99,999 94%). La calculele dispersiunii în cazul cercetărilor experimentale, care necesită o exactitate mai mare, calculăm, deci, deviația tip triplă; în cercetările clinice sau de teren ne mulțumim cu calcularea deviației-tip duble, care cuprinde, cum s-a arătat, 95,45% din totalul valorilor seriei de variație.

Corecția Sheppard. În cercetările referitoare, de exemplu, la înălțimea copiilor, nu putem merge la precizie deosebit de mare, fiind nevoiți,

de multe ori, să alegem intervale de 0,5 sau chiar 1 cm (la copii mai mari). În acest caz, mergînd la intervale de 1 cm, considerăm mijlocul intervalului cifra întreagă (de exemplu, 115, 116, 117 cm etc.). În intervalul corespunzător la 115 cm vom clasifica toate cazurile cu înălțimea cuprinsă între 114,6 și 115,5 cm, în intervalul următor, adică 116 cm vom clasifica toate cazurile cuprinse între 115,6 și 116,5 cm și așa mai departe. Pe bază de calcule s-a văzut că alegîndu-se, în mod practic, mijlocul intervalului se comite o eroare în plus în cadrul formulei deviației-tip, eroare care trebuie deci corectată. Se utilizează așa-zisa corecție a lui Sheppard, scăzîndu-se din formula deviației-tip o constantă egală cu $\frac{1}{12} = 0{,}083\,3$ din unitatea de interval utilizată.

Formula deviației-tip va deveni deci :

$$\sigma = \pm \sqrt{M_2 - (M_1)^2 - \frac{1}{12}} \text{ sau } \sigma = \pm \sqrt{M_2 - (M_1)^2 - 0{,}083\,3}$$

În exemplul anterior utilizînd corecția, vom obține :

$$\sigma = \pm \sqrt{1{,}7967} = \pm 1{,}34 \text{ cm.}$$

Vom avea, deci, deviația-tip egală cu $\pm$ 1,34 cm, în loc de $\pm$ 1,37 cm, cît ne-a dat formula fără corecția Sheppard.

Coeficientul de variabilitate. Deviația-tip (σ) exprimă dispersiunea de la medie în cifre absolute. În cazul înălțimii, σ se exprimă în cm sau mm ; la greutate, în grame sau multiplii și submultiplii lui ; în cazul mortalității infantile sau generale, în indici ; în cazul duratei medii de spitalizare, în zile etc.

Din aceste motive la colectivitățile caracterizate prin însușiri diferite, cum ar fi în dezvoltarea fizică : înălțimea și greutatea, deviațiile-tip, ale înălțimii medii și ale greutății medii nu sînt comparabile, deci nu putem cunoaște comparativ între ele variabilitățile celor două categorii de serii de valori, una exprimîndu-se în cm, iar cealaltă în grame.

Din aceste motive se utilizează coeficientul de variabilitate care nu este altceva decît deviația-tip, redată procentual.

Deci coeficientul de variație (cV) *este egal cu raportul dintre deviația-tip și media aritmetică ponderată* (Mp), formula fiind următoarea :

$$cV = \frac{\sigma \cdot 100}{Mp}$$

În cele două deviații-tip ale înălțimii la nou-născuții de sex masculin, coeficientul de variabilitate va fi* :

$$cV_1 = \pm \frac{1{,}60 \cdot 100}{50{,}64} = \pm 3{,}16\,\% ;$$

$$cV_2 = \frac{1{,}37 \cdot 100}{50{,}33} = \pm 2{,}72\,\%.$$

Se poate constata, cu ușurință, pe baza celor doi coeficienți de variabilitate, că dispersiunea celei de a doua colectivități este mai mică, deci eșantionul este mai omogen.

---

* Fără corecția Sheppard.

Un coeficient de variabilitate de ± 10% indică o dispersiune mică, deci o omogenitate a colectivității cercetate. Cînd coeficientul de variabilitate depășește ±10%, fiind situat între ±10 și ±20%, el indică o variabilitate medie, deci o colectivitate mai puțin omogenă. Peste ±20% coeficientul de variabilitate indică lipsa unei omogenități a colectivității, o dispersiune foarte mare a seriei de valori sau un număr prea scăzut de cazuri observate.

DEVIAȚIA TIP (σ) ÎN CAZUL SERIILOR DE VALORI CU FRECVENȚE EGALE

S-a arătat anterior că formula dezvoltată a deviației tip, conform momentului 1 (M_1) și momentului al 2-lea (M_2) este

$$\sigma = \pm \sqrt{\frac{\Sigma (X')^2 \cdot n}{N} - \frac{(\Sigma X' \cdot n)^2}{N^2}} = \pm \sqrt{M_2 - M_1^2}$$

În cazul în care se calculează deviația tip de la media unei serii de variație avînd frecvențe egale formula devine mult mai simplă, fără a mai fi nevoie de a mai recurge la calcularea momentului întîi și doi :

$$\sigma = \pm \sqrt{\frac{\Sigma (X')^2}{N}}$$

Să luăm un exemplu : tensiunea arterială luată la 8 persoane sănătoase, făcînd parte din aceeași grupă de vîrstă și sex a înregistrat următoarele valori : 110, 115, 116, 116, 119, 122, 124 și 130 mm Hg.

Media aritmetică simplă este de 119 mm Hg. Seria de valori ne arată însă o dispersiune importantă, o amplitudine de 20 mm Hg. (de la 110 mmHg la 130 mmHg).

În acest caz se pune pe drept întrebarea : este acest grup de persoane ales în mod corespunzător, deci lotul este omogen sau nu?

Pentru aceasta trebuie să se calculeze deviația tip și coeficientul de variație. Pentru deviația tip se scade media valorilor (119 mm Hg) din fiecare valoare a seriei, după care restul se ridică la pătrat. Suma pătratelor acestor diferențe (între valorile individuale și medie) se împarte la numărul de cazuri și apoi se extrage rădăcina pătrată :

Valorile (X)	Media aritmetică (Ma)	Diferența (deviația) (X')	Pătratul diferențelor (a deviațiilor) $(X')^2$
110	119	9	81
112	119	7	49
116	119	3	9
116	119	3	9
119	119	0	0
122	119	3	9
124	119	4	16
130	119	11	121

$$\Sigma (X')^2 = 294$$

$$\sigma = \pm \sqrt{\frac{\Sigma (X')^2}{N}} = \pm \sqrt{\frac{294}{8}} = \pm \sqrt{36,75} = \pm 6,06 \text{ mmHg}$$

$$cV = \pm \frac{\sigma \cdot 100}{Ma} = \pm \frac{606}{119} = \pm 5,09\%$$

Din calculul deviației tip și a coeficientului de variație rezultă că media tensiunii arteriale caracterizează bine lotul ales, care este omogen, cu un coeficient de variație scăzut de 5,09 %.

În cazul unui număr mic de observații, deci cu frecvențe mici (N mai mic de 30 de cazuri) se utilizează în calculul deviației tip, nu N ci N—1. Această scădere corectează o eroare de calcul ce se comite cu ocazia aplicării formulei deviației tip atunci cînd numărul de observații este mic (sub 30 de cazuri).

Formula, în această situație, devine :

$$\sigma = \pm \sqrt{\frac{\Sigma (X')^2}{N-1}} \rightleftharpoons \pm \sqrt{\frac{294}{8-1}} = \pm \sqrt{42,0} = \pm 6,48 \text{ mmHg.}$$

$$cV = \pm \frac{\sigma \cdot 100}{Ma} = \pm \frac{648}{119} = \pm 5,44\%.$$

8. EROAREA MEDIE A MEDIEI ARITMETICE

Media aritmetică calculată de noi nu reprezintă, de cele mai multe ori, decît o medie practică și nu una absolută, deoarece noi nu cercetăm totalitatea colectivității (totalitate denumită univers), ci doar un eșantion, o parte mai mare sau mai mică din total.

În cazul duratei medii de spitalizare în pneumonie sau în mortalitatea postoperatorie, ca și în dezvoltarea fizică la nou-născuți etc. noi nu cercetăm decît o mică parte din totalitatea colectivității. Din aceste motive s-ar putea ca eșantionul cercetat de noi să nu corespundă ca structură cu colectivitatea întreagă sau dacă alegerea a fost bine făcută, să existe, totuși, unele diferențe. În aceste condiții *media calculată diferă mai mult sau mai puțin de media reală, medie numită și absolută, care ar fi fost obținută în cazul cercetării totalității colectivității. Media practică prezintă, deci, o eroare față de această medie absolută*, eroare ce trebuie totdeauna calculată în cazul cercetărilor.

Indicele care măsoară eroarea valorii mediei este eroarea medie a mediei aritmetice ponderate (eM) care are următoarea formulă de calcul :

$$eM = \frac{\sigma}{\sqrt{N}}$$

Din formulă rezultă că eroarea medie a mediei aritmetice se calculează tot prin deviația-tip, căci în realitate *eroarea medie a mediei aritmetice corespunde, deviației-tip a unei serii întregi de medii, calculate fiecare pe un eșantion separat*, din cadrul aceleiași probleme.

Deci așa cum deviația-tip (σ) măsoară dispersiunea medie a valorilor unei colectivități față de media aritmetică ponderată, tot așa și *indicele erorii medii măsoară dispersiunea mediilor practice (obținute pe diferite eșantioane) față de media absolută, ideală.*

În practică, această eroare medie se calculează pornind de la media aritmetică ponderată a colectivității cercetate de noi.

Să calculăm eroarea medie la cele două eșantioane privind înălțimea la nou-născuți* :

$$eM_1 = \pm \frac{1,6}{\sqrt{258}} = \pm \frac{1,6}{16,1} = \pm 0,1 \text{ cm};$$

$$eM_2 = \pm \frac{1,37}{\sqrt{262}} = \pm \frac{1,37}{16,2} = \pm 0,08 \text{ cm}.$$

Din calculele de mai sus rezultă că eroarea medie a mediei aritmetice a primei colectivități este mai mare ca eroarea medie a celei de a doua colectivități, fiind de $\pm 0,1$ în primul caz și doar $\pm 0,08$ în al doilea caz.

Indicele erorii medii ne arată că *media absolută a înălțimii colectivității observate s-ar cuprinde între media practică (empirică) și plus sau minus eroarea ei medie.*

În cazul anterior, media absolută va fi cuprinsă între 50,64 $\pm$ 0,1 cm, adică între 50,54 și 50,74 cm. La colectivitatea a doua media absolută va fi cuprinsă între 50,33 $\pm$ 0,08 cm, adică între 50,25 și 50, 41 cm.

Cu alte cuvinte față de media absolută, media practică prezintă o eroare în medie de $\pm 0,1$ cm sau în al doilea caz de $\pm 0,08$ cm.

Ținînd seama de cele arătate la deviația-tip, pentru o mai mare exactitate, se ia eroarea medie dublă sau triplă, deoarece, în cazul erorii medii simple, doar în 68,26% din cazuri există probabilitatea ca media absolută să se cuprindă între media empirică și eroarea ei medie.

În cazul erorii medii duble sau triple probabilitatea ca media absolută să se găsească între medie și eroarea medie dublă sau triplă este de 95,45%, respectiv de 99,73%.

Deci pentru o exactitate mai mare vom lua eroarea medie triplă sau cel puțin dublă.

În exemplele anterioare vom obține :

$3\ eM_1 = \pm 0,1.3 = \pm 0,3.$

În cazul al doilea eroarea medie triplă va fi :

$3eM_2 = \pm 0,08.3 = \pm 0,24.$

Rezultă că media absolută se cuprinde cu o probabilitate de 99,73% din cazuri între medie $\pm$ 0,3 cm, respectiv $\pm 0,24$ cm, adică în primul caz în intervalul 50,34 — 50,94 cm, iar în al doilea caz în intervalul 50,09—50,57 cm.

După cum se poate vedea, deși cele două medii aritmetice ponderate au fost diferite (50,64 și 50,33 cm), erorile lor medii triple se întretaie.

Sugestiv, aceasta se poate constata din graficul din fig. 4.

Prima colectivitate are media aritmetică ponderată M, a doua colectivitate M'. Eroarea medie simplă, în primul caz, corespunde în grafic punctului $+a$ și $-a$; eroarea medie dublă punctului $+b$ și $-b$, iar eroarea medie triplă punctului $+c$ și $-c$.

La colectivitatea a doua eroarea medie simplă corespunde punctului $+a'$ și $-a'$, eroarea medie dublă punctului $+b'$ și $-b'$, iar eroarea medie triplă punctului $+c'$ și $-c'$.

---

* Fără corecție Sheppard.

Se poate constata din grafic că segmentele corespunzătoare punctelor $+c$ şi $-c'$, corespunzînd erorilor medii triple a celor două medii M şi M', se întretaie, de asemenea, şi punctele $+b$ şi $-b'$, deci erorile medii duble coincid pe o anumită porţiune.

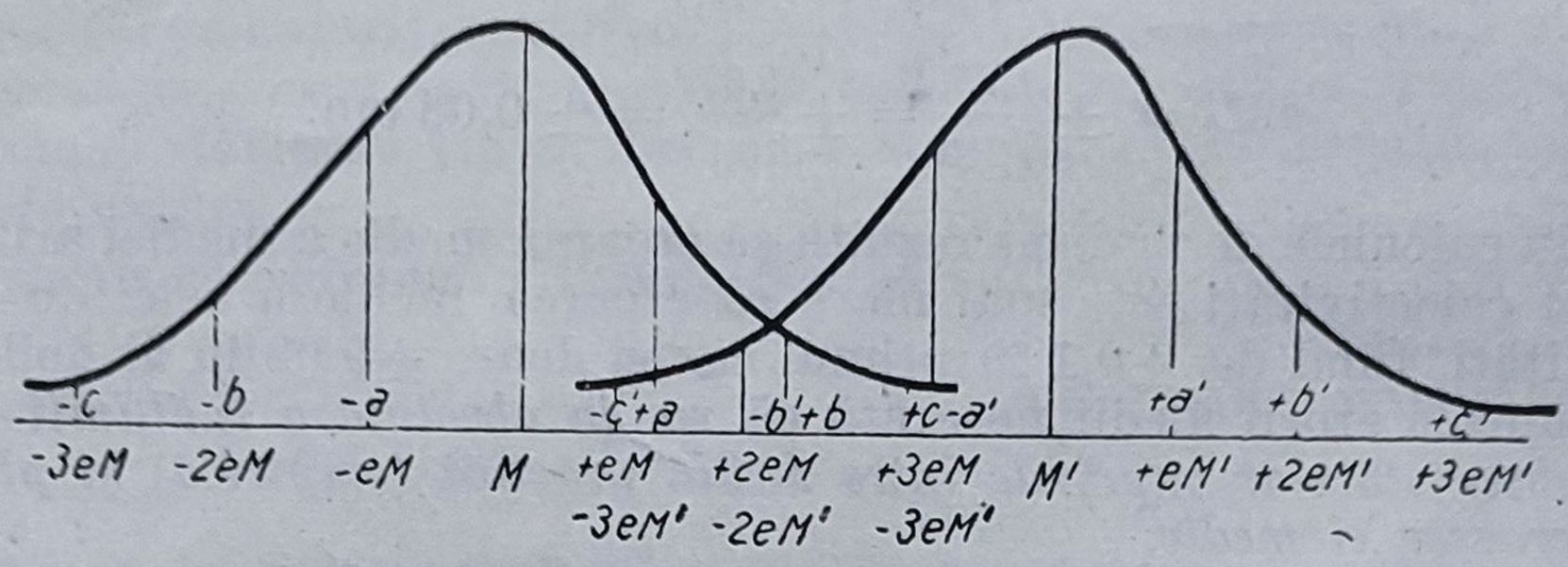

Fig. 4.

Aceasta înseamnă că diferenţa dintre cele două medii aritmetice nu este certă, cele două colectivităţi putînd avea o medie absolută sau ideală comună, care se găseşte între mediile aritmetice practice şi triplul erorii medii.

În cazul cînd cele două curbe nu s-ar fi întretăiat (în dreptul segmentelor ce corespund punctelor $+b$ şi $-b'$, sau $+c$ şi $-c'$), am fi putut afirma că diferenţa dintre cele două medii aritmetice ponderate este certă şi nu întîmplătoare. În această situaţie graficul ar fi luat aspectul din fig. 5.

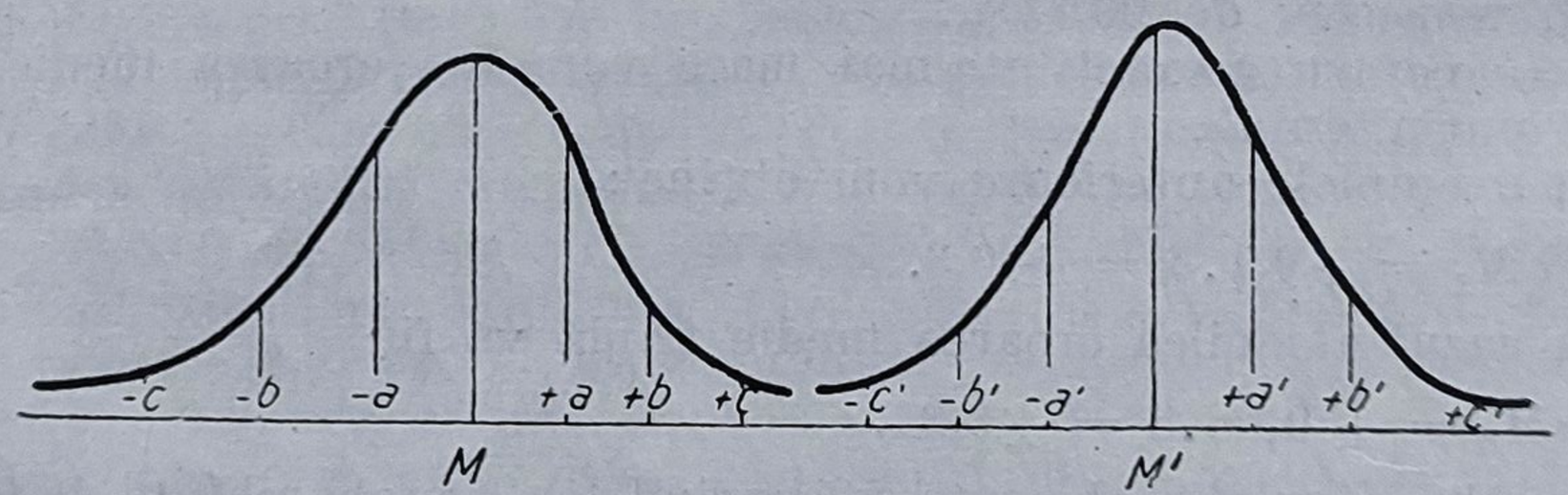

Fig. 5.

Eroarea diferenţei a două medii aritmetice. Cînd vrem să cunoaştem eroarea diferenţei între două medii, recurgem la un calcul special.

Formula de calcul al erorii medii a diferenţei dintre două medii aritmetice este următoarea :

eD (eroarea diferenţei dintre medii) $= \pm \sqrt{eM_1^2 + eM_2^2}$.

Deci *eroarea medie a diferenţei între cele două medii aritmetice ponderate (eD) va fi egală cu rădăcina pătrată a sumei celor două erori medii ridicate la pătrat.*

În exemplul nostru diferenţa dintre cele două medii este : 50,64 — 50,33 = 0,31. Rădăcina pătrată din suma pătratelor erorilor medii va fi :

$$eD = \pm \sqrt{eM_1^2 + eM_2^2} = \pm \sqrt{0{,}1^2 + 0{,}08^2} = \pm \sqrt{0{,}0164} = \pm 0{,}128$$

Luînd triplul diferenței erorilor medii obținem :

$$3.0{,}128 = \pm 0{,}384.$$

Comparînd diferența celor două medii aritmetice (0,31) cu triplul erorii medii, obținută mai sus ($\pm 0{,}384$), se poate constata, ca și în graficul din fig. 4, că triplul erorii medii a diferenței este mai mare decît diferența mediilor ponderate, care, deci, nu poate fi luată în considerare, diferența putîndu-se datori doar unei întîmplări.

Calcularea erorii medii a mediei aritmetice prezintă o importanță capitală, mai ales în medicina experimentală, unde diferențele în cadrul experimentării pe diferite eșantioane pot duce la interpretări juste sau nejuste, cu privire la eficacitatea unui medicament sau cu privire la doza letală a unor substanțe inoculate la animale de laborator, cu privire la eficacitatea unui vaccin etc.

Dată fiind importanța erorii medii, ea se trece totdeauna în paranteză, imediat după valoarea medie, de exemplu : 50,64 cm ($\pm$ 0,1) sau 50,33 cm ($\pm 0{,}08$).

Ca și în cazul deviației-tip (σ) și eroarea medie se calculează în cifre absolute : cm, g, l, indici de mortalitate etc. Pentru a putea compara eroarea medie la colectivități cu însușiri diferite (în cazul dezvoltării fizice : înălțimea, greutatea) se calculează procentual eroarea medie, după formula :

$$eM_1\% = \pm \frac{0{,}1 \cdot 100}{50{,}64} = \pm 0{,}20\%.$$

$$eM_2\% = \pm \frac{0{,}08 \cdot 100}{50{,}33} = \pm 0{,}16\%.$$

Un procent sub 1% al erorii medii a mediei aritmetice indică faptul că media aritmetică poate fi luată în considerare. Cu cît procentul va fi mai mic, cu atît probabilitatea va fi mai mare. Peste 3% eroarea medie denotă că media aritmetică ponderată nu poate fi luată în considerare, fiind incertă, ea rezultînd dintr-un eșantion nereprezentativ.

Din formula erorii medii se poate vedea că ea este direct proporțională cu mărimea deviației-tip (σ) și invers proporțională cu rădăcina pătrată a numărului total de cazuri cercetate (N). Deci *eroarea medie va fi cu atît mai mică, iar media aritmetică ponderată cu atît mai certă, cu cît numărul de observații va fi mai mare și cu cît dispersiunea colectivității va fi mai mică, adică atunci cînd eșantionul este mai omogen.*

Eroarea probabilă (*ep*) se poate utiliza în locul erorii medii a mediei aritmetice, existînd următoarea corespondență între ele :

$$ep = 0{,}6745\ eM \text{ și invers } eM = 1{,}4826\ ep.$$

Se vede, deci, că eroarea probabilă este mai mică decît eroarea medie a mediei aritmetice.

Semnificația erorii probabile este ca și cea a erorii medii, în unele lucrări și manuale utilizîndu-se eroarea probabilă în locul erorii medii.

Ea are un avantaj de calcul, prin faptul că între medie și $\pm ep$ se cuprind 50% din totalul valorilor seriei de variație în loc de 68,26% ca în cazul erorii medii a mediei aritmetice.

Eroarea probabilă dublă cuprinde 82,26% din aria curbei normale ; triplul erorii probabile 95,75%, iar cvadruplul erorii probabile cuprinde 99,30% din aria curbei normale.

9. MĂSURAREA DISPERSIUNII ŞI ERORII MEDII A FRECVENŢELOR RELATIVE. SEMNIFICAŢIA DIFERENŢEI DINTRE FRECVENŢELE RELATIVE

Frecvenţele relative, ca de exemplu frecvenţa letalităţii sau a vindecării, în cazul aplicării a două metode terapeutice într-o boală anumită, reprezintă de fapt medii aritmetice calculate pe un eşantion oarecare.

Dat fiind caracterul de variabilitate a reactivităţii organismului sănătos sau bolnav, frecvenţele relative de vindecare sau deces prezintă, de fapt, o medie de la care au loc o serie de deviaţii ale subgrupelor din colectivitatea cercetată. Aceste deviaţii iau, în general, forma distribuţiei normale, atunci cînd numărul de cazuri este mare şi destul de omogen.

Se aplică deci la frecvenţe calculul deviaţiei tip şi a erorii medii, precum şi calculul erorii medii a diferenţei.

Să luăm un exemplu de calcul al dispersiunii în cazul frecvenţelor relative :

Mortalitatea în pneumonie la un număr de 100 adulţi în vîrstă de 30—39 de ani, este de 3,0% în cazul unui tratament cu sulfamide şi doar de 1,5% în cazul tratamentului cu antibiotice la un lot egal de adulţi. Să vedem dacă diferenţa de 1,5% mortalitate denotă eficacitatea tratamentului cu antibiotice, sau se datoreşte doar unei întîmplări.

Dacă notăm la prima colectivitate mortalitatea cu P_1 şi frecvenţa vindecării cu Q_1, care ştim (de la calculul probabilităţii) că este egal cu $100-P_1$, iar la colectivitatea a doua aceleaşi frecvenţe le notăm cu P_2 şi Q_2 egal cu $100-P_2$, deviaţia tip va fi următoarea :

$$\varepsilon\ \sigma_1 = \pm\sqrt{\frac{P_1 \cdot Q_1}{N_1}} = \pm\sqrt{\frac{P_1 \cdot (100-P_1)}{N_1}};$$

$$\varepsilon\ \sigma_2 = \pm\sqrt{\frac{P_2 \cdot Q_2}{N_2}} = \pm\sqrt{\frac{P_2 \cdot (100-P_2)}{N_2}}.$$

Înlocuind cu cifre datele din formulă vom obţine :

$$\sigma_1 = \pm\sqrt{\frac{3{,}0 \cdot (100-3{,}0)}{100}} = \pm\sqrt{\frac{3{,}0 \cdot 97{,}0}{100}} = \pm\sqrt{2{,}91} = \pm 1{,}71;$$

$$\sigma_2 = \pm\sqrt{\frac{1{,}5 \cdot (100-1{,}5)}{100}} = \pm\sqrt{\frac{1{,}5 \cdot 98{,}5}{100}} = \pm\sqrt{1{,}48} = \pm 1{,}22.$$

Erorile medii se calculează conform formulei :

$$eM_1 = \pm\frac{\sigma_1}{\sqrt{N_1}} \text{ şi } eM_2 = \pm\frac{\sigma_2}{\sqrt{N_2}};$$

$$eM_1 = \pm\frac{1{,}71}{\sqrt{100}} = \pm\frac{1{,}71}{10{,}00} = \pm 0{,}17;$$

$$eM_2 = \pm\frac{1{,}22}{\sqrt{100}} = \pm\frac{1{,}22}{10{,}00} = \pm 0{,}12.$$

Pentru a putea deduce dacă diferența dintre cele două medii ale frecvențelor de mortalitate este certă sau nu, se va calcula eroarea diferențelor între cele două frecvențe : 3,0% și 1,5%.

Diferența între medii este de 3,0%—1,5% = 1,5%.

Eroarea diferenței se calculează după formula arătată anterior :

$$eD = \pm\sqrt{eM_1^2 + eM_2^2} = \pm\sqrt{0{,}17^2 + 0{,}12^2} = \pm\sqrt{0{,}03 + 0{,}01} =$$

$$= \pm\sqrt{0{,}04} = \pm 0{,}2.$$

Deci, mortalității constatate la cele două colectivități, care au o diferență a frecvențelor de 1,5% îi corespunde o eroare medie a diferențelor de ± 0,2, care înmulțită cu trei (triplul erorii medii) dă ± 0,6, valoare mai mică decît diferența celor două letalități 1,5.

Rezultă că diferența mortalității la cele două colectivități nu este întîmplătoare și că tratamentul cu antibiotice a dat, în mod cert și nu întîmplător, o mortalitate mai mică.

În cazul frecvențelor mai mici de 200, formula anterioară a deviației tip : $\sigma = \pm\sqrt{\frac{P \cdot Q}{N}}$ trebuie corectată, deoarece această formulă aplicîndu-se perfect în frecvențele teoretice, dă rezultate diferite în cazul unui N mic și o probabilitate (P) aproape de 50,0%. Aceste corective se expun în paragraful următor.

10. MĂSURAREA ERORII MEDII A MEDIEI ARITMETICE ÎN CAZUL FRECVENȚELOR MICI

În cazul frecvențelor mai mici de 200, măsurarea erorii medii a mediei aritmetice ponderate nu ne dă valori exacte, căci eroarea mediei este cu atît mai mare, cu cît frecvența este mai mică.

Sub limita de 200 de observații formula erorii medii (eM) nu se poate utiliza, decît cu o probabilitate mică.

În aceste cazuri se utilizează tabele speciale de distribuție Student, prin care putem să obținem valorile erorii medii și în cazul cifrelor mici, utilizînd valoarea t, conform tabelului VII.

Tabelul VII

Valoarea t în măsurarea erorii medii atunci cînd frecvența (n) este mică

Frecvența (n)	t	n	t	n	t	n	t
5	6,6	11	4,0	17	3,5	50	3,2
6	5,5	12	3,9	18	3,5	60	3,1
7	4,9	13	3,8	19	3,5	100	3,1
8	4,5	14	3,7	20	3,4	200	3,0
9	4,3	15	3,6	30	3,3		
10	4,1	16	3,6	40	3,2		

Să luăm un exemplu privind mortalitatea în operațiile de rezecție gastrică în boala ulceroasă, pe un număr de 50 de cazuri, indicele de mortalitate fiind de 6%.

Deviația-tip în cazul acestei frecvențe relative va fi :

$$\sigma = \pm \frac{\sqrt{6 \cdot (100 - 6)}}{50} = \pm \sqrt{11{,}28} = \pm 3{,}36$$

Eroarea medie conform formulei $eM = \pm \frac{\sigma}{\sqrt{n}}$ va fi :

$$eM = \pm \frac{3{,}36}{\sqrt{50}} = \pm \frac{3{,}36}{7{,}07} = \pm 0{,}48\%.$$

Știm că pentru o probabilitate cît mai mare (99,73%) se ia triplul erorii medii. Triplul erorii medii va fi : 3. ($\pm$ 0,48%) = $\pm$ 1,44%. Aceasta înseamnă că limitele dispersiunii valorii medii (6,0%) față de media absolută vor fi cuprinse cu o probabilitate de 99,73% între 6,0% $\pm$ 1,44, adică între 4,56% și 7,44%.

Ținînd seama de numărul scăzut de observații (50 de cazuri) triplarea erorii medii se face utilizînd tabelele de distribuție Student. Deci în loc să triplăm eroarea medie a mediei aritmetice (în cazul nostru $\pm$ 0,48) căutăm din tabelul VII valoarea *t* ce corespunde unei frecvențe (*n*) de 50 de cazuri. În tabel, aceasta este de 3,2. Înmulțind eroarea medie cu 3,2, rezultatul va fi $\pm 0{,}48 . 3{,}2 = \pm 1{,}54\%$. Deci media mortalității în operațiile de rezecție gastrică pentru ulcer va fi de 6,0% ($\pm$1,54), avînd, deci, o variabilitate cuprinsă între limitele : 6,0% $\pm$ 1,54, adică între 4,46% și 7,54%.

11. MĂSURAREA DISPERSIUNII ÎN CAZUL EVENIMENTELOR RARE „LEGEA NUMERELOR MICI"

Atunci cînd probabilitatea ca un eveniment oarecare să se producă este foarte mică sau cînd la un număr oarecare de observații este chiar zero, deși știm că la un număr mai mare de observații acest eveniment ar putea să survină, calculul dispersiunii și al erorii medii a frecvențelor relative, expus anterior, nu mai este valabil.

În aceste cazuri se utilizează alte metode pentru calcularea probabilității de apariție a evenimentelor rare, sau a acelor evenimente care nu apar din cauza numărului mic de observații.

În cazul probabilităților de apariție a evenimentelor rare (*p*, fiind foarte mic) se utilizează în locul distribuției normale, distribuția Poisson, distribuție numită încă și „legea numerelor mici".

Graficul din figura 6 ne arată diferite distribuții Poisson cu *m* crescînd. Se poate constata că la un *m* (probabilitate de apariție a evenimentului) mic, distribuția este unilaterală. O dată cu creșterea lui *m* distribuția devine din ce în ce mai puțin asimetrică, pentru ca la un *m* egal cu 10 să se transforme într-o distribuție normală.

În cazul unui eveniment egal cu zero, cînd există, totuși, certitudinea că este posibilă apariția acestui eveniment la un număr mai mare de observații, se utilizează graficul din fig. 7. La un număr *n* de cazuri observate se citește, în partea dreaptă a coloanei, p_o, *n*, care corespunde

diviziunilor lui n şi ne arată limita superioară a probabilităţii de apariţie a evenimentului.

Formula de calcul a cunoaşterii limitei superioare a probabilităţii de apariţie a fenomenului în caz că ea a fost zero (p_o) este următoarea :

$$m = \frac{p_{0,n}}{n}$$

Să luăm un exemplu : mortalitatea la un număr de 40 de bolnavi de difterie a fost zero, cu toate că ştim că ar fi putut surveni o mortalitate

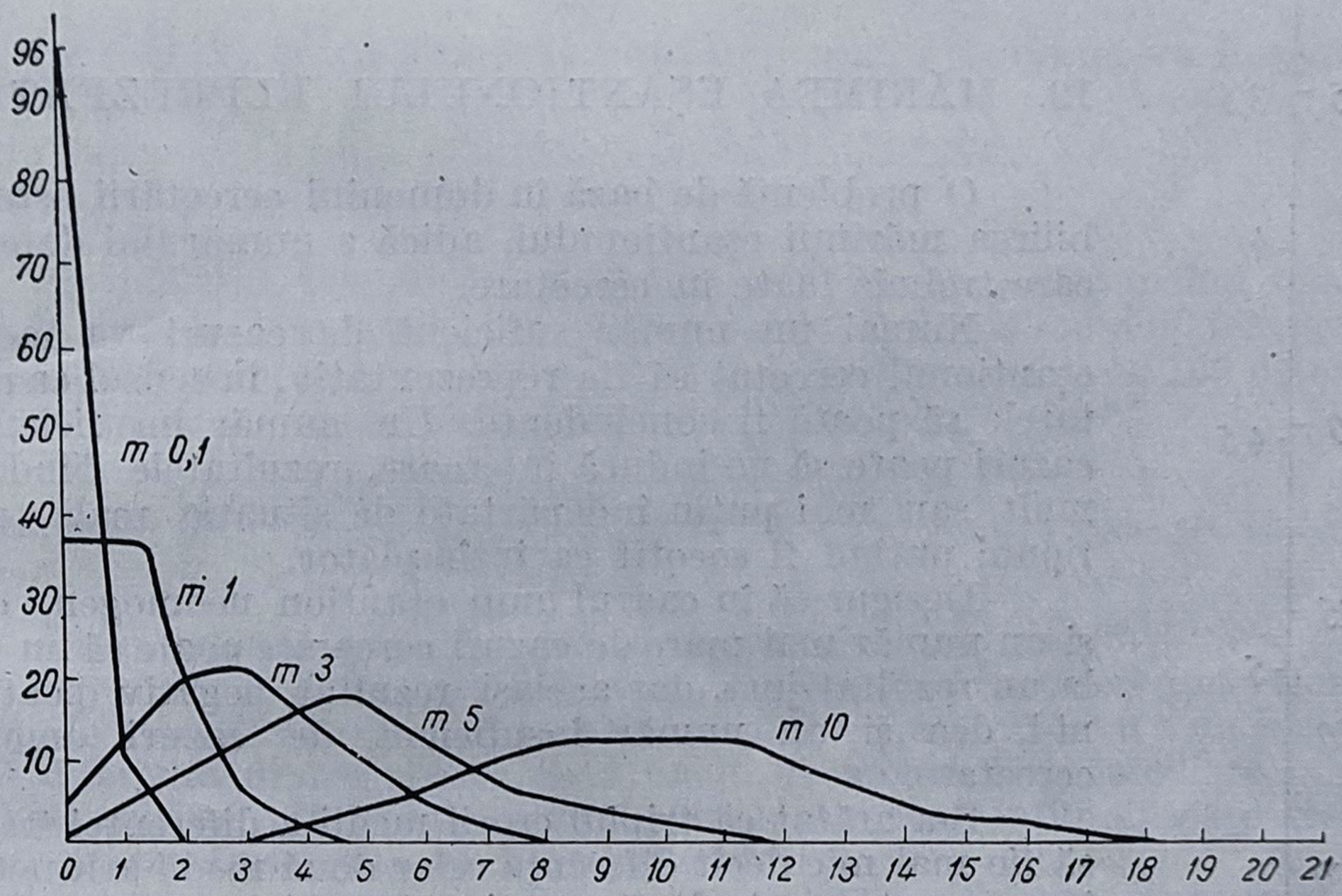

Fig. 6. — Distribuţia Poisson.

oarecare, observînd un număr mai mare de cazuri. Care ar putea fi totuşi, în cazul nostru, limita superioară a mortalităţii ?

Citim pe diagramă în dreptul lui $n = 40$ cît revine lui $p_{o.n}$ şi aflăm valoarea 5,5. Avînd elementele cifrice, completăm formula anterioară şi vom găsi că limita superioară probabilă a mortalităţii ar putea fi de 13,8%.

$$m = \frac{5,5}{40} = 0,138\,;\; m\% = 0,138.100 = 13,8\%.$$

Deci cu o probabilitate de 99,73% (corespunzînd lui 3σ) se poate afirma, în cazul nostru, că mortalitatea în difterie nu ar fi depăşit limita de 13,8% în cazul observării unui număr mai mare de bolnavi.

Aceeaşi metodă se utilizează cînd de exemplu $p = n$, deci cînd probabilitatea de a se produce un eveniment a fost egală cu numărul de cazuri observate.

Să luăm un exemplu. Dintr-un număr de 10 bolnavi de tetanos toate cazurile au decedat, deşi ştim că există cazuri de vindecare. Care ar fi

fost limita superioară a mortalității la un număr suficient de cazuri observate?

Citim de pe diagramă $p_{o,n}$ corespunzător unui n egal cu 10. Acesta este de 4,46. Aplicăm formula și apoi transformăm în procente.

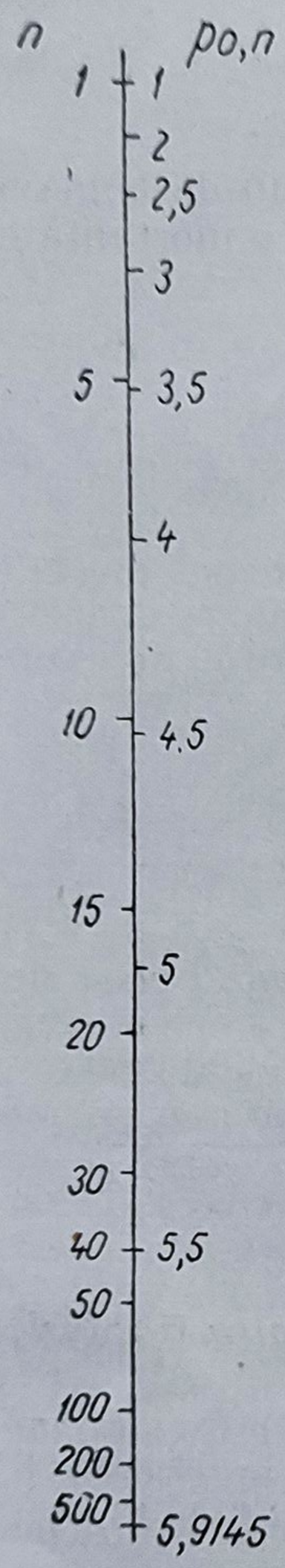

Fig. 7. — Diagrama pentru calculul limitei superioare a probabilității apariției evenimentului cînd p = 0.

$$m = \frac{[p_{o,n}}{n} = \frac{4,46]}{10} = 0,446\,;\ m\% = 44,6\%.$$

Deci la un număr mai mare de cazuri cu o probabilitate de 99,73% am putea susține că limita superioară a mortalității n-ar fi depășit 44,6% din cazuri.

12. MĂRIMEA EȘANTIONULUI REPREZENTATIV

O problemă de bază în domeniul cercetării este stabilirea mărimii eșantionului, adică a numărului de cazuri care trebuie luate în cercetare.

Numai un număr suficient de cazuri va face ca eșantionul cercetat să fie reprezentativ, în sensul ca rezultatele să poată fi concludente. Un număr insuficient de cazuri poate să ne inducă în eroare, rezultatele fiind mai mult sau mai puțin îndepărtate de situația reală, eșantionul putînd fi socotit ca întîmplător.

Desigur că în cadrul unui eșantion neomogen, chiar și un număr mai mare de cazuri cercetate poate să nu ducă la un rezultat just, dar același rezultat negativ poate să ni-l dea și un număr insuficient de cazuri omogene cercetate.

S-a arătat că triplul erorii medii a diferenței trebuie să fie mai mic decît diferența celor două medii aritmetice; în caz contrar cele două eșantioane nu sînt reprezentative.

Dar în cercetările clinice, de exemplu, nu stă în puterea noastră să mărim numărul de observații.

În acest caz este necesar să se asocieze mai mulți cercetători din diverse spitale pentru ca din observațiile comune, bazate pe aceeași metodologie, să se poată trage concluzii juste.

Cîte cazuri trebuie încă observate atunci cînd diferența erorii medii indică un eșantion insuficient?

Pentru a rezolva această problemă pornim tot de la calculul dispersiunii arătat anterior în cazul frecvențelor relative:

$$\sigma_1 = \pm\sqrt{\frac{P_1 \cdot (100 - P_1)}{N_1}} \text{ și } \sigma_2 = \pm\sqrt{\frac{P_2 \cdot (100 - P_2)}{N_2}}$$

Dacă vom calcula dispersiunea diferenței (σD) vom obține, ca și la calculul erorii medii a diferenței:

$$\sigma D = \pm\sqrt{\sigma_1^2 + \sigma_2^2}$$

Dacă înlocuim σ_1^2 și σ_2^2 din formulă cu corespondentul lor din formula anterioară obținem :

$$\sigma D = \pm \sqrt{\frac{P_1 \cdot (100 - P_1)}{N_1} + \frac{P_2 \cdot (100 - P_2)}{N_2}}.$$

Atunci cînd $N_1 = N_2$, adică cele două eșantioane sînt egale, formula de mai sus va fi :

$$\sigma D = \pm \sqrt{\frac{P_1 \cdot (100 - P_1) + P_2 \cdot (100 - P_2)}{N}}.$$

De aici N, adică numărul total de cazuri observate, va fi egal cu :

$$N = \frac{P_1 \cdot (100 - P_1) + P_2 \cdot (100 - P_2)}{\sigma D^2}.$$

Să înlocuim formulele anterioare prin cifrele corespunzînd mortalității în bronhopneumonie, de exemplu, după două metode terapeutice (sulfamide și antibiotice), mortalitatea la primul eșantion de 100 de cazuri fiind de 8,0 %, iar la al doilea eșantion de 100 de cazuri 1,0 %.

$$\sigma D = \pm \sqrt{\frac{8 \cdot (100 - 8) + 1 \cdot (100 - 1)}{100}} = \sqrt{\frac{736 + 99}{100}} = \sqrt{8,35} = \pm 2,9.$$

Dacă luăm triplul lui σD obținem $\pm 8,7$, adică ($3. \pm 2,9 = \pm 8,7$), deci dispersiunea diferenței este mult peste valoarea diferenței celor doi indici de mortalitate care este : 8,0 % — 1,0 % = 7,0 %. Pentru a ne convinge dacă diferența dintre aceste două medii de mortalitate calculată la suta de bolnavi din spitale este semnificativă sau doar întîmplătoare trebuie ca eșantioanele folosite să fie reprezentative.

Pentru ca eșantionul să fie reprezentativ ar trebui să creștem numărul de observații (N), pînă cînd deviația tip a diferenței (σD) scade la o valoare mult mai mică, astfel ca triplul ei să nu fie mai mare de 7,0 (diferența dintre cele două mortalități), adică să nu fie mai mare decît $\pm 2,33$ (deci triplul ei $3.2,33 = 6,99$).

În acest caz formula eșantionului reprezentativ (N) ar fi :

$$N = \frac{8 \cdot (100 - 8) + 1 \cdot (100 - 1)}{2,33^2} = \frac{835}{5,43} = 154.$$

Deci numărul cazurilor luate în cercetare ar trebui să fie 154. În acest caz am putea afirma cu probabilitate de 99,73 % (3σ) că eșantioanele sînt reprezentative.

Dacă ne-am mulțumi cu o probabilitate de 95,45 % (2σ), eșantionul reprezentativ ar trebui să fie mai mic, și anume 68 bolnavi în loc de 154 (în acest caz σD va fi de 3,5, care luat de două ori dă 7,0, deci atît cît este diferența dintre cei doi indici de mortalitate) :

$$N = \frac{8 \cdot (100 - 8) + 1 \cdot (100 - 1)}{3,5^2} = \frac{835}{12,25} = 68$$

13. MĂSURAREA DIFERENȚEI DINTRE SERIILE DE VARIAȚIE PRIN METODA χ^2 (CHI PĂTRAT)

Măsurătorile mediei nu ne arată asemănarea sau diferența între două sau mai multe serii de variație fără a face întreg șirul de calcule ale mediei, ale dispersiunii de la medie și ale erorii medii.

Altfel nu ne putem da seama dacă două serii, cu aceleași valori medii, sînt sau nu asemănătoare ca structură, dacă există o diferență esențială între ele sau dacă diferențele sînt doar întîmplătoare.

Procedeul χ^2 simplifică mult măsurarea diferenței între seriile de variație, fără a mai fi nevoie de a calcula indicii de dispersiune :

Metoda χ^2 se aplică îndeosebi în acele situații cînd evenimentele așteptate se exclud unul pe celălalt, în sensul că nu este posibil să se producă decît unul dintre ele, ca de exemplu : deces sau nu în cazul unor boli, complicații sau nu în diferitele operații etc.

Să luăm un exemplu din domeniul medicinei experimentale :

La un grup de 30 de animale de laborator acțiunea unei substanțe (A) s-a soldat cu un număr de 18 decese, deci o mortalitate de 60%. Aceeași substanță A în combinație cu o altă substanță B, la un alt lot alcătuit tot din 30 de animale a cauzat decesul la 10 animale, deci o mortalitate de 33,3% din cazuri.

Problema se pune dacă diferențele de mortalitate sînt esențiale, datorită substanței B ce s-a adăugat în plus primei substanțe A sau dacă diferențele sînt pur întîmplătoare.

Pentru această problemă se efectuează calculul χ^2 (chi pătrat), pe baza tabelei cvadruple. Se utilizează numai cifrele absolute, nu cele relative (tabelul VIII).

Tabelul VIII

Metoda χ^2 în măsurarea diferenței dintre 2 serii

	Au decedat	Nu au decedat	Total
Primul lot supus substanței A	18_1 (a_1)	12_1 (b_1)	30 ($a_1+b_1=n_1$)
Al doilea lot supus substanței A + B	10_2 (a_2)	20_2 (b_2)	30 ($a_2+b_2=n_2$)
TOTAL	28 ($a_1+a_2=n_3$)	32 ($b_1+b_2=n_4$)	60 ($n_1+n_2=n_3+n_4=N$)

χ^2 se calculează în felul următor :

$$\chi^2 = \frac{(a_1 \cdot b_2 - b_1 \cdot a_2)^2}{n_1 \cdot n_2 \cdot n_3 \cdot n_4} \cdot (N - 1);$$

$$\chi^2 = \frac{(18 \cdot 20 - 12 \cdot 10)^2}{30 \cdot 30 \cdot 28 \cdot 32} \cdot (60 - 1) = 4{,}21.$$

A. Balan.

În cazul în care χ^2 calculat este mai mare decît 9, înseamnă că diferența dintre cele două serii comparate nu este întîmplătoare. Această afirmație o putem face cu probabilitate de 99,73%.

Cînd ne mulțumim cu o probabilitate de 99,0%, χ^2 trebuie să fie mai mare decît 6,64, iar în cazul unei probabilități de 95,0%, χ^2 trebuie să fie mai mare de 3,84.

În cazul nostru diferența de letalitate este demonstrată cu o probabilitate de peste 95,0%, deoarece χ^2 este 4,21.

Specificăm că măsurarea diferenței dintre serii de variație prin metoda χ^2 se poate efectua doar în cazul valorilor care să depășească 5 sau şi mai bine 10 cazuri.

Metoda χ^2 se aplică şi în cazul măsurării diferenței mai multor serii.

14. CORELAȚIA DINTRE FENOMENE. REGRESIUNEA

Corelația. Cînd vrem să cunoaștem dacă există sau nu vreo legătură de dependență între seriile de variabilitate a două sau mai multe fenomene, recurgem la calculul corelațiilor.

Corelația dintre seriile de variabilitate dintre două fenomene trebuie să fie bilaterală, în sensul că la orice variabilitate a primului fenomen să-i corespundă o variabilitate a celui de-al doilea fenomen.

Un exemplu de corelație este cea dintre înălțimea şi greutatea corpului. Cu cît înălțimea este mai mare, cu atît, în general, şi greutatea este mai mare şi invers.

În acest caz, cînd variabilitatea între cele două fenomene are loc în același sens, vorbim de o corelație pozitivă. În cazul în care variabilitatea dintre cele două fenomene are loc în sens contrar, vorbim de o corelație negativă.

Corelația se poate constata sugestiv din tabelul de corelație (vezi tabelul IX).

Din tabelul IX se poate constata că pentru fiecare înălțime a copiilor îi corespunde o anumită greutate şi invers, pentru fiecare greutate îi corespunde o anumită înălțime. Se constată, de exemplu, din tabelul de corelație, că dintre cei 6 copii de 10 ani cu înălțimea de 126 cm, 3 au greutatea de 26 kg, iar cîte unul are greutatea de 24, 25 şi 27 kg.

De asemenea, dintre cei 9 copii cu greutatea de 25 kg, 4 au înălțimea de 128 cm, 3 înălțimea de 127 cm şi cîte unul are înălțimea de 126 cm, respectiv 129 cm.

Din tabelul IX se poate observa că este cazul a două serii de variație, şi anume înălțimea şi greutatea la 425 de copii de 10 ani de sex masculin, prezentate nu separat, ci în dependență, în corelație una cu cealaltă.

Greutățile prezentate în coloanele verticale se notează cu x, iar înălțimile prezentate pe rîndurile orizontale se notează cu y. Ultima coloană verticală cuprinde suma lui y, totalul cazurilor de o anumită înălțime, iar ultimul rînd — orizontal — cuprinde suma lui *x*, adică totalul cazurilor de o anumită greutate.

Tabelul IX

Corelaţia între înălţimea şi greutatea corpului la copiii de 10 ani de sex masculin

y = înălţimea în cm \ x = greutatea în kg	24	25	26	27	28	29	30	31	32	33	34	35	Suma y
126	1	1	3	1									6
127		3	4	1		1							9
128	1	4	5	7	1	2	1						21
129		1	2	6	9	6	4	1					29
130			1	7	36	17	6	2	1				70
131			1	6	27	56	39	18	4	1			152
132				2	10	16	26	7	2	1			64
133					1	7	10	11	6	2	1		38
134						1	2	4	7	3	1		18
135							1	2	3	1	1		8
136									1	2	2	1	6
137										1	2	1	4
Suma x	2	9	16	30	84	106	89	45	24	11	7	2	425

Cele două totaluri generale sînt egale (în cazul nostru 425), deoarece fiecare caz a fost situat într-un pătrat care corespunde unei anumite înălţimi şi greutăţi.

Se mai poate constata din tabelul IX că această corelaţie este cu atît mai puternică cu cît dispersiunea cazurilor de la valoarea medie este mai mică şi, invers : cu cît dispersiunea de la medie este mai mare, cu atît corelaţia este mai slabă.

Care este gradul de corelaţie nu-l putem cunoaşte decît prin calculul coeficientului de corelaţie (notat cu *r*).

Coeficientul de corelaţie (*r*) se calculează din tabelul de corelaţie.

În primul rînd se calculează deviaţia de la medie a ambelor serii de variaţie, cea a greutăţii şi cea a înălţimii copiilor, ca medie ipotetică fiind aleasă — ca şi în cazul calculării momentelor — coloana şi rîndul cu frecvenţa cea mai mare, care se încadrează în linii mai groase. Se obţin astfel patru pătrate (tabelul X) corespunzînd următoarelor corelaţii :

— primul, pătratul superior stîng, corelaţiei cazurilor cu greutatea sub valoarea medie şi înălţimea tot sub valoarea medie ;

— al doilea, pătratul superior drept, corelaţiei cazurilor cu greutatea peste valoarea medie şi înălţimea sub valoarea medie ;

— al treilea, pătratul inferior stîng, corelaţiei cazurilor cu greutatea sub valoarea medie şi înălţimea peste valoarea medie ;

— al patrulea, pătratul inferior drept, corelaţiei cazurilor cu greutatea peste valoarea medie şi înălţimea tot peste valoarea medie.

Tabelul X

Corelaţia între înălţimea şi greutatea corpului la copiii de 10 ani de sex masculin

y = înălţimea în cm		x = greutatea în kg											
		24	25	26	27	28	29	30	31	32	33	34	35
		−5	−4	−3	−2	−1	0	+1	+2	+3	+4	+5	+6
126	−5	1 (25)	1 (20)	3 (15)	1 (10)								
127	−4		3 (16)	4 (12)	1 (8)		1 (0)						
128	−3	1 (15)	4 (12)	5 (9)	7 (6)	1 (3)	2 (0)	1 (3)					
129	−2		1 (8)	2 (6)	6 (4)	9 (2)	6 (0)	4 (2)	1 (4)				
130	−1			1 (3)	7 (2)	36 (1)	17 (0)	6 (1)	2 (2)	1 (3)			
131	0			1 (0)	6 (0)	27 (0)	56 (0)	39 (0)	18 (0)	4 (0)	1 (0)		
132	+1				2 (2)	10 (1)	16 (0)	26 (1)	7 (2)	2 (3)	1 (4)		
133	+2					1 (2)	7 (0)	10 (2)	11 (4)	6 (6)	2 (8)	1 (10)	
134	+3						1 (0)	2 (3)	4 (6)	7 (9)	3 (12)	1 (15)	
135	+4							1 (4)	2 (8)	3 (12)	1 (16)	1 (20)	
136	+5									1 (15)	2 (20)	2 (25)	1 (30)
137	+6										1 (24)	2 (30)	1 (36)

În tabelul X în fiecare pătrat mic se găsesc două valori, prima notată în stînga, partea de sus a pătratului reprezentînd cazurile care se găsesc în tabelul IX, iar a doua valoare în dreapta şi în partea de jos a pătratului corespunzînd produsului dintre abaterile de la valoarea medie a greutăţii şi înălţimii notate cu plus şi minus în dreptul fiecărui rînd şi coloană.

De exemplu, în primul pătrăţel din stînga şi de sus prima valoare (1) corespunde cazului care prezintă înălţimea de 126 cm şi greutatea de 24 kg, iar valoarea din josul şi dreapta primului pătrăţel (25) corespunde produsului deviaţiilor de la media înălţimii şi greutăţii : $(-5)\cdot(-5) = 25$. La fel se procedează şi în continuare cu fiecare pătrăţel. În pătrăţelul care corespunde valorii medii ipotetice, cuprinse între liniile groase, deviaţia de la medie este zero.

După semnele algebrice se poate vedea că primul pătrat (stînga sus) şi ultimul (dreapta jos) vor avea rezultatul produsului cu semnul +, iar al doilea pătrat (dreapta sus) şi al treilea (stînga jos) vor avea rezultatul produsului notat cu semnul —.

Se trece în etapa următoare la înmulţirea celor două valori din fiecare pătrăţel şi apoi la totalizarea sumei acestor produse, ţinînd seama, bineînţeles, de semnele algebrice.

Făcînd suma celor patru pătrate obţinem : + 472 — 28 — 16 + + 667 = + 1 095.

Acest rezultat (1 095) ne arată deviaţia celor două însuşiri (înălţimea şi greutatea) de la media aritmetică arbitrar aleasă, înmulţite cu frecvenţa respectivă.

Tabelul X

Corelația între înălțimea și greutatea corpului la copiii de 10 ani de sex masculin

y = înălțimea în cm		x = greutatea în kg											
		24	25	26	27	28	29	30	31	32	33	34	35
		−5	−4	−3	−2	−1	0	+1	+2	+3	+4	+5	+6
126	−5	1 25	1 20	3 15	1 10								
127	−4		3 16	4 12	1 8		1 0						
128	−3	1 15	4 12	5 9	7 6	1 3	2 0	1 3					
129	−2		1 8	2 6	6 4	9 2	6 0	4 2	1 4				
130	−1			1 3	7 2	36 1	17 0	6 1	2 2	1 3			
131	0			1 0	6 0	27 0	56 0	39 0	18 0	4 0	1 0		
132	+1				2 2	10 1	16 0	26 1	7 2	2 3	1 4		
133	+2					1 2	7 0	10 2	11 4	6 6	2 8	1 10	
134	+3						1 0	2 3	4 6	7 9	3 12	1 15	
135	+4							1 4	2 8	3 12	1 16	1 20	
136	+5									1 15	2 20	2 25	1 30
137	+6										1 24	2 30	1 36

În tabelul X în fiecare pătrat mic se găsesc două valori, prima notată în stînga, partea de sus a pătratului reprezentînd cazurile care se găsesc în tabelul IX, iar a doua valoare în dreapta și în partea de jos a pătratului corespunzînd produsului dintre abaterile de la valoarea medie a greutății și înălțimii notate cu plus și minus în dreptul fiecărui rînd și coloană.

De exemplu, în primul pătrățel din stînga și de sus prima valoare (1) corespunde cazului care prezintă înălțimea de 126 cm și greutatea de 24 kg, iar valoarea din josul și dreapta primului pătrățel (25) corespunde produsului deviațiilor de la media înălțimii și greutății : $(-5)\cdot(-5) = 25$. La fel se procedează și în continuare cu fiecare pătrățel. În pătrățelul care corespunde valorii medii ipotetice, cuprinse între liniile groase, deviația de la medie este zero.

După semnele algebrice se poate vedea că primul pătrat (stînga sus) și ultimul (dreapta jos) vor avea rezultatul produsului cu semnul +, iar al doilea pătrat (dreapta sus) și al treilea (stînga jos) vor avea rezultatul produsului notat cu semnul —.

Se trece în etapa următoare la înmulțirea celor două valori din fiecare pătrățel și apoi la totalizarea sumei acestor produse, ținînd seama, bineînțeles, de semnele algebrice.

Făcînd suma celor patru pătrate obținem : $+472 - 28 - 16 + + 667 = +1\,095$.

Acest rezultat (1 095) ne arată deviația celor două însușiri (înălțimea și greutatea) de la media aritmetică arbitrar aleasă, înmulțite cu frecvența respectivă.

Primul pătrat (+)	Al doilea pătrat (−)	Alt treilea pătrat (−)	Al patrulea pătrat (+)
1 . 25 = 25	2 . 6 = 12	1 . 3 = 3	2 · 2 = 4
1 . 20 = 20	6 . 4 = 24	4 . 2 = 8	10 · 1 = 10
3 . 15 = 45	9 · 2 = 18	1 · 4 = 4	1 · 2 = 2
1 · 10 = 10	1 · 3 = 3	6 · 1 = 6	− 16
3 · 16 = 48	7 · 2 = 14	2 · 2 = 4	
4 · 12 = 48	36 · 1 = 36	1 · 3 = 3	
1 · 8 = 8	+472	− 28	
1 · 15 = 15			
4 · 12 = 48			
5 · 9 = 45			
7 · 6 = 42			
1 · 3 = 3			
1 · 8 = 8			

26 · 1 = 26	1 · 15 = 15
7 · 2 = 14	1 · 4 = 4
2 · 3 = 6	2 · 8 = 16
1 · 4 = 4	3 · 12 = 36
10 · 2 = 20	1 · 16 = 16
11 · 4 = 44	1 · 20 = 20
6 · 6 = 36	1 · 15 = 15
2 · 8 = 16	2 · 20 = 40
1 · 10 = 10	2 · 25 = 50
2 · 3 = 6	1 · 30 = 30
4 · 6 = 24	1 · 24 = 24
7 · 9 = 63	2 · 30 = 60
3 · 12 = 36	1 · 36 = 36
	+ 667

Pentru a afla media aritmetică ponderată se calculează momentul 1 al înălțimii și greutății, al căror produs se adună sau se scade din rezultatul nostru (1 095).

Momentul 1 al seriei de variație a înălțimii, calculat după regulile arătate la calculul dispersiunii de la valoarea medie este de 0,06, iar momentul 1 al seriei de variație a greutății este de 0,24. Produsul lor dă 0,01. Acesta adunat la 1 095 dă 1 095,01.

Acest rezultat se împarte apoi la produsul dintre deviațiile tip al celor două serii de variație : greutate (x) și înălțime (y), înmulțit cu numărul total de cazuri (N) conform formulei coeficientului de corelație :

$$r = \frac{\Sigma\, n \cdot x' y'}{N \cdot \sigma x \cdot \sigma y},$$

în care :

r = coeficientul de corelație ;
n = frecvența, numărul de cazuri din fiecare pătrățel ;
x' = deviația de la medie a intervalelor de greutate ;
y' = deviația de la medie a intervalelor de înălțime ;
Σ = suma, adică totalul general al produselor : $n.\ x'.\ y'$;
N = numărul total de cazuri ;
σx = deviația tip de la media ponderată a greutății ;
σy = deviația tip de la media ponderată a înălțimii.

Înlocuind în formula de mai sus simbolurile cu cifrele calculate obținem :

$$r = \frac{1\,095{,}01}{N \cdot \sigma x \cdot \sigma y}$$

$$N = 425$$

$$\sigma x = \pm 1{,}83$$

$$\sigma y = \pm 1{,}86$$

$$r = \frac{1\,095{,}01}{1\,445{,}00} = \pm 0{,}76.$$

Coeficientul de corelaţie se cuprinde între + 1 şi —1. Cu cît valoarea coeficientului se apropie de + 1 sau — 1, cu atît corelaţia este mai accentuată, fie în acelaşi sens (+) fie în sens opus (—).

Din contră, cu cît valoarea coeficientului se apropie de zero, cu atît corelaţia între cele două serii de variaţie este mai slabă, putînd arăta că probabil nici nu există vreo corelaţie între cele două însuşiri la care s-a calculat coeficientul r; se poate deci deduce că nu există nici o corelaţie cînd r este egal cu zero.

Cînd r are valoarea cuprinsă între $\pm$ 0,5 şi $\pm$ 0,7 vorbim de o corelaţie mijlocie, peste $\pm$ 0,7 este cazul unei corelaţii accentuate, iar cînd r este egal cu $\pm$ 1 obţinem o corelaţie perfectă.

Corelaţia se poate reprezenta şi grafic (vezi graficul din fig. 8). Cu cît corelaţia va fi mai accentuată, cu atît relieful graficului va fi mai accentuat şi mai puţin extins în lăţime.

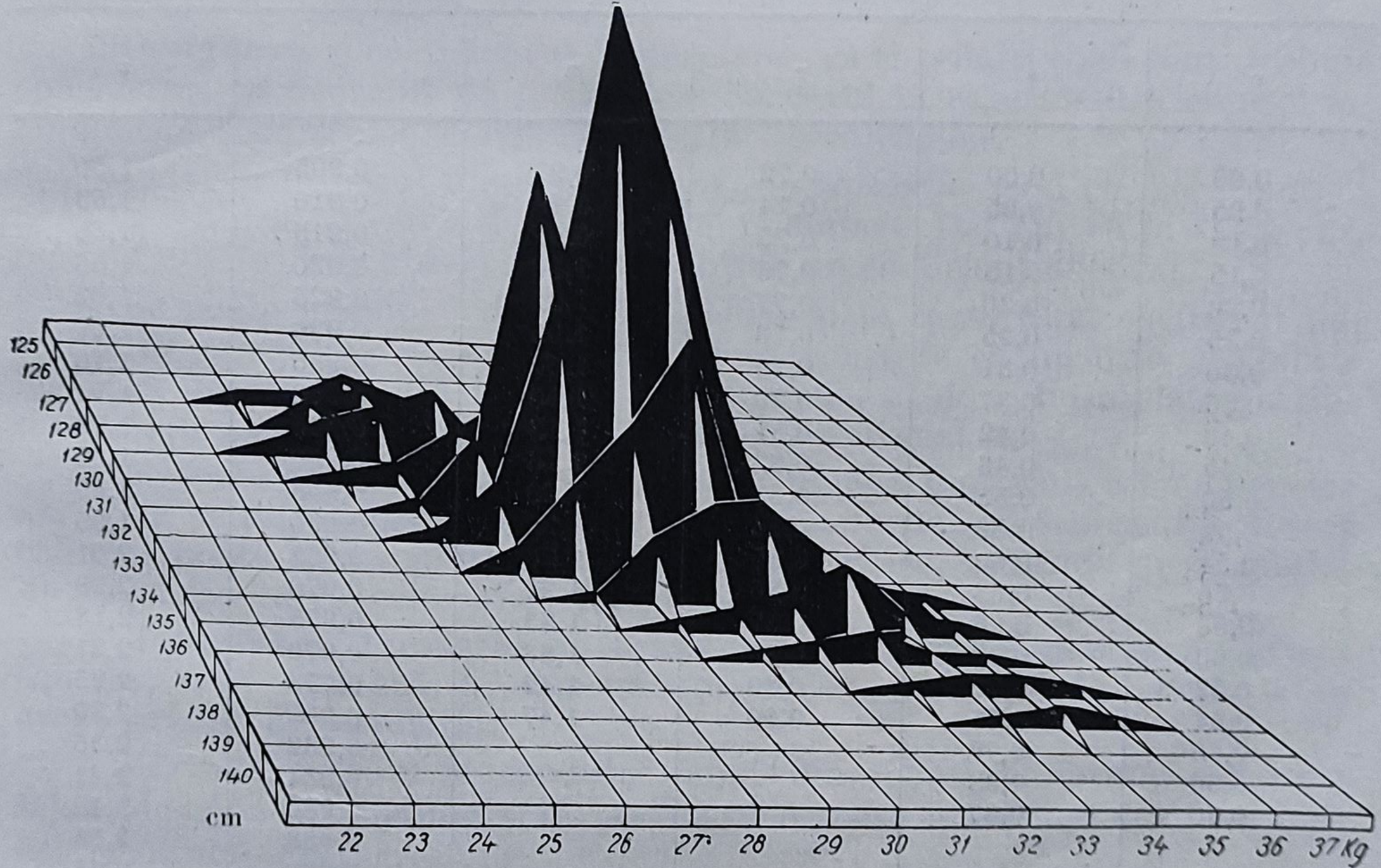

Fig. 8. — Grafic de corelaţie.

Pentru o mai mare siguranţă a coeficientului de corelaţie se calculează, şi în acest caz, eroarea lui medie, care va fi cu atît mai mică, cu cît numărul de cazuri omogene observate va fi mai mare, deci ca şi în cazul erorii medii a mediei aritmetice.

Calculul se efectuează pe baza formulei următoare:

$$er = \pm \frac{1 - r^2}{\sqrt{N}}.$$

În cazul nostru, înlocuind simbolurile cu cifre, vom obţine:

$$er = \pm \frac{1 - 0{,}76^2}{\sqrt{425}} = \pm \frac{1 - 0{,}58}{20{,}62} = \pm 0{,}02.$$

Deci coeficientul de corelaţie este cuprins între 0,76 ± 0,02, adică între 0,74 şi 0,78, sau luînd triplul erorii medii : (3) · (± 0,02) = ± 0,06 se va afla între 0,70 şi 0,82.

Pe baza erorii medii se poate constata (ca şi în cazul valorilor medii), dacă există sau nu diferenţe esenţiale între doi coeficienţi de corelaţie.

La un număr mic de observaţii eroarea medie a coeficientului *r* nu este aplicabilă. În acest caz se utilizează un alt parametru (*z*), pe baza unor tabele ale lui Fisher, unde cunoscînd coeficientul *r* putem citi şi coeficientul *z*, cu care putem să procedăm la calculul erorii medii şi în cazul unui număr mic de observaţii (tabelul XI).

Tabelul XI

Cifre de corelaţie

r	*z*	*r*	*z*	*r*	*z*
0,00	0,00	0,72	0,91	0,905	1,50
0,05	0,05	0,74	0,95	0,910	1,53
0.10	0,10	0,75	0,97	0,915	1,56
0,15	0,15	0,76	1,00	0,920	1,56
0,20	0,20	0,77	1,02	0,925	1,62
0.25	0,25	0,78	1,05	0,930	1,66
0,30	0,31	0,79	1,07	0,935	1,70
0,35	0,37	0,80	1,10	0,940	1,74
0,40	0,42	0,81	1,13	0,945	1,78
0,45	0,48	0,82	1,16	0,950	1,83
0,50	0,55	0,83	1,19	0,955	1,89
0,52	0,58	0,84	1,22	0,960	1,95
0,54	0,60	0,85	1,26	0,965	2,01
0,56	0,63	0,86	1,29	0,970	2,09
0,58	0,66	0,87	1,33	0,975	2,18
0,60	0,69	0,88	1,38	0,976	2,21
0,62	0,72	0,89	1,42	0,978	2,25
0,64	0,76	0,90	1,47	0,980	2,30
0,66	0,79			0,982	2,35
0,68	0,83			0,984	2,41
0,70	0,87			0,986	2,48
				0,988	2,55
				0,990	2,65
				0,992	2,76
				0,994	2,90
				0,996	3,11
				0,998	3,46
				1,000	∞

Dacă în exemplul anterior, coeficientul de corelaţie (*r*) ar fi fost de 0,77 şi ar fi corespuns unui număr scăzut de observaţii, 28 de cazuri în loc de 425, în tabelul XI ar corespunde un coeficient *z* egal cu 1,02.

Din acest tabel rezultă că în cazul unui coeficient de corelaţie (*r*) egal cu zero sau foarte mic, indicînd lipsa unei corelaţii, îi corespunde un coeficient *z* egal cu zero sau foarte mic ; dar coeficientul *z* creşte mai repede, odată cu creşterea coeficientului *r*, depăşind 1,0, în timp ce coeficientul maxim de corelaţie nu poate depăşi ±1,0.

În cazul unui număr mic de observații, în locul erorii medii a coeficientului de corelație (r) se utilizează eroarea medie a coeficientului z, care se calculează după formula următoare :

$$\sigma z = \frac{1}{\sqrt{n-3}} .$$

Completînd simbolurile din formulă cu cifrele din exemplul anterior, vom obține :

$$\sigma z = \frac{1}{\sqrt{28-3}} = \frac{1}{5} = 0{,}20$$

Aceasta înseamnă că valoarea lui z este cuprinsă între 0,82 și 1,22 (1,02 $\pm$ 0,20 = 0,82 pînă la 1,22).

Desigur că și coeficientul de corelație, ca și ceilalți coeficienți trebuie considerați ca deosebit de utili, doar în cazul în care analiza esenței fenomenelor studiate concordă cu rezultatele calculelor.

Chiar și în cazul unui coeficient de corelație egal cu zero putem avea în realitate o corelație între două fenomene. Desigur că în acest caz, mărind numărul de observații, vom obține un coeficient adecvat.

Se poate întîmpla, din contră (chiar și în cazul unui coeficient puternic de corelație), ca analiza critică să constate că nu este posibilă o legătură de dependență între două fenomene, ci doar întîmplător seriile de variație s-au potrivit, dînd un coeficient apropiat de ± 1.

Cu toate acestea, atunci cînd analiza esenței materiale a fenomenelor ne arată o interdependență între două fenomene, coeficientul r corect calculat poate întări presupunerea analizei critice, arătîndu-ne și gradul de corelație între cele două fenomene.

Regresiunea. Corelația ne arată dacă între două fenomene există sau nu vreo legătură de dependență, iar coeficientul de corelație ne arată gradul de corelație.

Mărimea cu care se modifică unul din fenomene, atunci cînd celălalt fenomen se modifică și invers, nu ne-o poate arăta însă decît calculul coeficientului de regresiune.

Coeficientul de regresiune ne arată, deci, mărimea cu care crește sau scade un element, atunci cînd celălalt element cu care este în legătură de dependență crește sau scade cu o unitate de măsură.

Ca și în cazul corelației, și regresiunea poate avea același sens, adică la o creștere a unui element să crească și celălalt element sau să aibă sens opus, adică în cazul creșterii unui element să se micșoreze celălalt. Deci regresiunea poate fi pozitivă sau negativă.

În exemplul anterior, privind copiii de 10 ani de sex masculin, calculînd coeficientul de regresiune pentru greutate, vom putea constata cu cît se modifică greutatea atunci cînd înălțimea crește sau scade cu 1 cm. În cazul înălțimii, coeficientul de regresiune va arăta cît crește înălțimea cînd greutatea crește cu 1 kg.

Formulele de calcul pentru coeficienţii de regresiune, notînd însuşirile cu x (greutatea în cazul nostru) şi cu y (înălţimea), vor fi următoarele :

$$R_{\frac{x}{y}} = r \cdot \frac{\sigma x}{\sigma y};$$

$$R_{\frac{y}{x}} = r \cdot \frac{\sigma y}{\sigma x}.$$

În primul caz vom cunoaşte cu cît se modifică greutatea, cînd înălţimea creşte cu 1 cm sau invers cu cît scade greutatea cînd înălţimea scade cu 1 cm ; deci vom obţine regresiunea lui x faţă de y.

În al doilea caz, vom avea regresiunea lui y (înălţimea) faţă de x (greutatea).

Înlocuind simbolurile cu cifrele deja cunoscute anterior pentru r, σ x şi σ y vom avea :

$$R_{\frac{x}{y}} = 0{,}76 \cdot \frac{1{,}83}{1{,}86} = 0{,}74 \text{ kg};$$

$$R_{\frac{y}{x}} = 0{,}76 \cdot \frac{1{,}86}{1{,}83} = 0{,}78 \text{ cm}$$

Primul coeficient de regresiune $R_{\frac{x}{y}}$ ne arată că la o modificare cu 1 cm a înălţimii corespunde o modificare a greutăţii de 0,74 kg în medie.

Al doilea coeficient de regresiune $R_{\frac{y}{x}}$ ne arată că unei modificări cu 1 kg a greutăţii îi corespunde o modificare a înălţimii de 0,78 cm în medie.

Se pot stabili corelaţii şi regresiuni nu numai între două fenomene, ci şi mai multe. În acest caz se alcătuiesc tabele de corelaţie pentru fiecare pereche de elemente.

De exemplu în cazul a trei însuşiri : x, y, z, vom alcătui 3 tabele de corelaţie : x şi y, x şi z, y şi z.

Pentru patru însuşiri : x, y, z, şi a vom avea 6 tabele de corelaţie : x şi y, x şi z, x şi a, y şi z, y şi a, z şi a.

Pentru cinci însuşiri vom avea 10 tabele de corelaţie etc.

La fel se poate stabili şi regresiunea a mai multor însuşiri.

Corelaţia şi regresiunea se poate stabili în numeroase probleme ale medicinei clinice şi experimentale ca, de exemplu : corelaţia între temperatură şi bătăile cardiace, între durata graviditătii şi greutatea şi înălţimea nou-născutului, precocitatea internării, aplicării tratamentului şi mortalitatea în abdomen acut, pneumonie, bronhopneumonie etc.

CAPITOLUL AL II-LEA

METODE MATEMATICE ÎN CERCETĂRILE MEDICALE EXPERIMENTALE

1. INTRODUCERE

Metodele matematice sînt relativ frecvent utilizate în medicina experimentală, dată fiind problematica ce necesită utilizarea acestei metode în fundamentarea ştiinţifică a cercetărilor.

Complexitatea fenomenelor biologice şi variabilitatea lor, datorită factorilor multipli ce le determină, impun — pe lîngă metoda observaţiei — utilizarea şi a metodei experimentale, pentru a putea ajunge la deducţii logice, cu privire la cauzele care provoacă apariţia acestor fenomene, evoluţia, dispariţia sau menţinerea lor.

Stabilirea legăturii complexe de la cauză la efect are o însemnătate de bază, atît teoretică, cît şi practică. Aceasta este însă uneori greu de realizat, atît din cauza complexităţii factorilor determinanţi şi adjuvanţi ce produc sau influenţează diversele fenomene biologice, cît mai ales din cauza complexităţii şi variabilităţii cu care se produce reactivitatea organismelor vii, la diversele influenţe externe sau interne.

Din aceste considerente este necesar să se utilizeze, cît mai mult posibil, împreună metoda experimentală cu cea a observaţiei, pentru a putea determina, în mod ştiinţific, legăturile de cauzalitate dintre diferitele fenomene.

Metodele matematice ajută pe cercetător doar să întărească constatările făcute pe calea observaţiei şi a experimentului, prin analiza esenţei materiale a fenomenelor.

Metodologia statisticii matematice în sine nu poate, deci, să constituie baza deducţiilor pentru cercetarea dată ; ea nu poate înlocui deducţiile logice ale observaţiei şi ale experimentului.

Dacă metodele matematice nu pot fi considerate singure ca baza deducţiilor logice ale observaţiei şi ale experimentelor, nu este mai puţin adevărat că utilizarea lor este de o deosebită importanţă.

Astfel stabilirea grupărilor diferitelor elemente, a valorilor medii şi a erorilor acestor valori, a gradului de dispersiune a cazurilor individuale faţă de valoarea medie, a certitudinii diferenţei între două rezultate de aceeaşi natură, sau între fenomenele care apar la un grup de animale

de laborator și grupul martor în cazul unor experimente, a numărului necesar de animale de laborator în diferite cercetări etc. sînt absolut necesare.

Din aceste considerente metodele matematice trebuie să fie utilizate în toate observațiile și experimentele biologice, în cazul nostru în medicina experimentală, pentru ca deducțiile logice, obținute prin analiza și sinteza științifică a fenomenelor, să aibă fundamentarea necesară.

Observația și cercetarea experimentală în domeniul medicinei duce la o diversitate de rezultate chiar și în cazul în care condițiile de experimentare sînt identice.

Aceasta se datorește faptului că, spre deosebire de experimentarea din domeniul chimiei sau fizicii, unde putem reproduce condiții în general asemănătoare, în domeniul biologiei aceste condiții identice sau asemănătoare nu se pot reproduce în aceeași măsură, dată fiind complexitatea fenomenelor biologice, diversitatea reactivității organismelor vii etc.

Astfel dacă vom cerceta metabolismul bazal la 10 sau 20 de persoane de același sex, vîrstă, greutate, înălțime etc. (acestea fiind perfect sănătoase), cu aceeași aparatură și în aceleași condiții prealabile, vom obține — și în acest caz — o diversitate de rezultate, chiar dacă rezultatele sînt dispersate normal în jurul unei medii.

Tot astfel în cazul animalelor de laborator, dacă vom alege 10 șobolani pentru a avea un grup absolut omogen, în vederea inoculării cu o anumită substanță medicamentoasă, vom constata că rezultatele, fie în ceea ce privește mortalitatea, fie chiar în ceea ce privește concentrația substanței în urina șobolanilor vor fi diferite, deși în general apropiate de o valoare medie.

Față de valorile medii vom avea, și în primul și în al doilea rînd, cazuri care se depărtează mai mult, iar altele, din contra, mai puțin.

Variabilitatea cu care răspunde organismul uman (chiar în cazul unei selecționări absolut omogene) este și mai mare în cazul anumitor boli.

Menționăm, însă, că o omogenitate absolută, în cazul experimentelor mici nu este necesară, deoarece, în acest caz, n-am mai reproduce situația reală, care se caracterizează tocmai printr-o mare variabilitate.

De altfel, problema omogenității se pune în felul următor: în ce măsură lotul experimental este reprezentativ pentru întreaga colectivitate de indivizi din care a fost ales, în ce măsură concluziile obținute pe acest lot sînt valabile pentru colectivitate. Din aceste considerente o selecționare de indivizi de o omogenitate oricît de perfectă, dar nereprezentativă, ne-ar putea duce, de cele mai multe ori, la rezultate inutilizabile pentru restul colectivității.

Deci selecționarea cazurilor utilizate în vederea cercetărilor experimentale trebuie să țină seama de omogenitate, fără a pierde însă din vedere ca lotul cercetat să fie reprezentativ pentru colectivitatea la care căutăm să aplicăm deducțiile cercetării.

Aplicarea metodelor matematice devine deci indispensabilă în aprecierea exactității rezultatelor obținute în urma cercetărilor, deoarece chiar deducțiile logice devin, de cele mai multe ori, incerte în cazul nefolosirii acestor metode.

De asemenea, în cazul deducţiilor nelogice, metodologia statistică ne ajută să descoperim erorile ce au survenit în concluziile observaţiilor şi experimentelor.

În cele ce urmează vom reda cîteva din metodele matematice cele mai uzuale întrebuinţate în cercetările experimentale, rămînînd ca pentru adîncirea unor probleme să se recurgă la capitolul I, noţiuni generale.

2. LOTUL EXPERIMENTAL ŞI LOTUL MARTOR

Cînd se urmăreşte, spre exemplu, cunoaşterea influenţei pe care o au anumite substanţe medicamentoase sau vaccinuri asupra organismului uman sau se urmăreşte o metodă nouă de diagnostic sau tratament se utilizează metoda observaţiei şi a experimentului.

De la caz la caz, după natura problemei, experimentele se efectuează pe diferite categorii de animale de laborator sau chiar pe oameni.

Alegerea subiecţilor, fie animale, fie oameni, depinde de o serie de factori pe care cercetătorul trebuie să-i determine, în funcţie de scopul urmărit.

Această alegere nu poate deci să fie făcută spontan, ci cu ajutorul analizei ştiinţifice a problemei ce se cercetează, analiză făcută pe bază de documentare.

Desigur că cercetarea se poate baza şi pe o serie de ipoteze ce urmează a fi verificate prin experimentare, cu condiţia ca aceste ipoteze să aibă o fundamentare ştiinţifică.

Această alegere cade, deci, în sarcina cercetătorului, fără să se poată fixa de la început reguli privind categoria de indivizi utilizaţi, mai ales atunci cînd nu s-au mai făcut cercetări identice sau asemănătoare.

Problema se complică atunci cînd se trece la lotul de subiecţi ce se alege pentru experimentare şi la lotul martor.

În această problemă cercetătorii pot greşi de multe ori, atunci cînd nu cunosc anumite reguli în ceea ce priveşte mărimea şi numărul eşantionului care formează lotul respectiv de subiecţi sau indivizi.

Cîte animale sau cîte persoane trebuie luate în lotul experimental şi cîte în lotul martor? Din cîte grupuri trebuie alese aceste eşantioane? Cît de omogene trebuie să fie aceste loturi?

Iată probleme la care cercetătorul trebuie să răspundă de la început, după ce scopul şi ipoteza de lucru a fost stabilită, iar fără utilizarea metodelor matematice răspunsul va fi de multe ori eronat.

Problema numărului de eşantioane sau din cîte grupe de indivizi trebuie alese subiectele depinde de scopul cercetării.

În cazul persoanelor gruparea se va face pe sexe, grupe de vîrstă, condiţii social-economice şi cultural-sanitare, condiţii de dezvoltare psihofizică etc.

În cazul animalelor, gruparea va fi mai simplă : specie, sex, vîrstă, mărime etc.

Cînd presupunem că o anumită substanţă sau un anumit procedeu va putea influenţa în mod diferit rezultatele scontate, vom alege eşantioane separate din diferite grupe, pentru a obţine loturi cît mai tipice.

Cînd vrem să diferenţiem rezultatele obţinute, cunoscînd că există deosebiri importante în ceea ce priveşte reactivitatea diferitelor grupe de

indivizi, trebuie să utilizăm, de asemenea, loturi separate din aceste grupe (pe sex, vîrstă etc.).

Mai ales atunci cînd se stabileşte doza maximă sau doza letală la un anumit medicament va trebui să se acorde cea mai mare atenţie în alegerea lotului, care trebuie să fie omogen, dar şi reprezentativ pentru întreaga colectivitate.

O atenţie deosebită trebuie acordată şi alegerii lotului martor. Acest lot martor este necesar în multe cercetări experimentale, deoarece cu ajutorul lui putem elimina riscul unor variaţii întîmplătoare a lotului experimental.

S-ar putea concepe, uneori, lipsa lotului martor, utilizînd doar lotul experimental, la care să observăm apariţia sau dispariţia unor fenomene în legătură cu substanţele sau procedurile utilizate de experimentator, după ce, în prealabil, am studiat acest lot înainte de a se fi aplicat experimentul.

Ţinînd seama de faptul că există posibilitatea de variabilitate în timp (zilnică, sezonieră sau periodică), va fi greu de stabilit ulterior în ce măsură diferenţele în modificările din cadrul lotului experimental, înainte şi după aplicarea experimentului, se datoresc acestui experiment şi în ce măsură variabilităţii în timp. Din aceste motive se utilizează, de cele mai multe ori, şi lotul martor.

Lotul martor va fi ales pe baza aceloraşi considerente ca şi lotul experimental, adică ţinînd seamă de scopul urmărit, de ipoteza de lucru şi de factorii de omogenitate şi reprezentativitate.

În orice caz lotul martor trebuie să fie ales din aceeaşi grupare ca şi lotul experimental, pentru a se putea elimina deosebirile în reactivitatea lotului, datorită nu experimentului, ci alegerii neomogene a acestuia.

În cadrul experimentelor cu animale de laborator de multe ori lotul experimental se sacrifică sau suferă o serie de modificări care îl fac inapt pentru repetarea, în aceleaşi condiţii, a experimentului. În această situaţie lotul trebuie schimbat cu un lot nou, ales în aceleaşi condiţii.

Lotul martor nu trebuie însă reînnoit totdeauna, deoarece, în general, el nu suferă modificări.

Atunci cînd experimentul este mai complex, în sensul că lotului experimental i se administrează, de exemplu, o substanţă toxică A şi lotului martor două substanţe, A şi B, ultima anihilînd acţiunea toxică a substanţei A, s-ar putea ca şi în lotul martor să aibă loc modificări care să necesite sacrificarea acestuia sau în cazul în care nu este necesară sacrificarea, să nu poată fi, totuşi, utilizabil în alte experimente.

În general, se merge pe economisirea maximă a numărului de animale de laborator şi a substanţelor utilizate, fără ca aceasta să se facă în detrimentul eşantionului reprezentativ, necesar în cercetările experimentale.

Care este mărimea eşantionului, deci numărul de indivizi din lotul experimental şi lotul martor? Răspunsul depinde de următoarele probleme ce trebuie cunoscute, o dată cu fixarea scopului şi a ipotezei de lucru:

— Dacă pe acelaşi lot experimental şi lot martor se efectuează un experiment cronic, repetat pe etape, în aceleaşi condiţii, cu aceleaşi substanţe, această repetare compensează un număr mai mare de indivizi. Deci, cu cît se efectuează mai multe experimente şi în aceleaşi condiţii pe unul şi acelaşi lot experimental, cu atît numărul de indivizi din acest lot poate fi mai scăzut.

— Cu cît modificările în lotul experimental care au loc cu ocazia experimentelor sînt mai accentuate, mai vizibile, sau mai de lungă durată și aceste modificări se repetă cu aceeași intensitate la fiecare experiment succesiv, efectuat bineînțeles în aceleași condiții, cu atît numărul de indivizi din lotul experimental și lotul martor poate fi, în general, mai mic. Din contră, dacă modificările ce au loc sînt mai puțin vizibile, mai puțin intense, sau dacă la repetare sînt uneori rezultate contradictorii, numărul de indivizi din lotul experimental sau lotul martor trebuie să fie mai mare.

— Eșantionul martor și experimental poate fi cu atît mai mic, cu cît limitele de variabilitate a fenomenului cercetat sînt mai înguste.

— Cînd lotul experimental sau lotul martor este mai omogen, eșantionul poate fi mai mic, atît la lotul experimental, cît și la lotul martor.

— Cu cît experimentul este în general mai simplu, înțelegînd prin aceasta evidențierea unei legături mai nete între cauză și efect, cu atît lotul poate fi mai mic.

Deci esențialul în stabilirea mărimii eșantionului este de a elimina cît mai mult posibilitatea modificărilor întîmplătoare și de a obține o cît mai mare certitudine în stabilirea legăturilor de cauzalitate ce au loc cu ocazia experimentelor.

Pe baza acestor reguli generale se va putea trece la fixarea lotului experimental și a lotului martor. Pe baza metodelor matematice se poate stabili cu exactitate mărimea eșantionului reprezentativ, deci dacă lotul ales a fost sau nu suficient ca mărime și ca număr (deci ca grupare omogenă și reprezentativă).

Specificăm din nou că analiza calitativă a problemei ce se experimentează trebuie considerată ca determinantă în elucidarea faptului dacă rezultatele obținute sînt sau nu posibile, sau doar întîmplătoare. Metodele matematice întăresc doar justețea ipotezelor și a concluziilor obținute.

3. REZULTATUL MEDIU, DISPERSIUNEA, CERTITUDINEA DIFERENȚEI A DOUĂ MEDII. CALCULUL EȘANTIONULUI REPREZENTATIV

Avînd în vedere complexitatea fenomenelor biologice, rezultatele cercetărilor experimentale prezintă o variabilitate sensibilă.

Există, cu toate acestea, o tendință centrală, în sensul că la un număr mai mare de subiecte sau la repetarea de mai multe ori a experimentelor, rezultatele obținute se grupează în jurul unei valori medii, medie de la care există o dispersiune mai mică sau mai mare a cazurilor individuale, în funcție de omogenitatea lotului ales, cît și de variabilitatea fenomenelor cercetate.

Numărul mic de subiecte sau indivizi utilizați uneori cu ocazia experimentului face ca legea numerelor mari să nu mai poată fi aplicată în stabilirea indicilor medii și a dispersiunii de la medie, ca în cazul eșantioanelor mai mari (peste 30 de indivizi dintr-o grupare).

Să luăm un exemplu de calculare a indicilor amintiți : experimentînd acțiunea antagonistă a două substanțe A și B, dintre care prima toxică, iar a doua ce anihilează acțiunea primei substanțe, pe un lot expe-

rimental şi altul martor, obţinem următorul rezultat : din lotul experimental, compus din 60 de şobolani, căruia i s-a administrat intraperitoneal o anumită doză din substanţa A, după un interval de 6 ore, au pierit 33 de şobolani. Din lotul martor, compus din 30 de şobolani, cărora li s-a administrat în condiţii similare substanţa A, dar concomitent şi substanţa B, pier, în acelaşi interval de 6 ore, 6 şobolani.

Ţinînd seama de faptul că în ambele loturi au avut loc decese, deşi este evident că în cazul lotului experimental proporţia este cu mult mai mare şi ţinînd seamă şi de numărul diferit de şobolani din cele două loturi, se pune problema dacă se poate considera ca certă acţiunea toxică a substanţei A şi acţiunea antitoxică, antagonistă a substanţei B.

De asemenea, se pune problema de a determina mortalitatea la şobolani, datorită substanţei toxice A şi mortalitatea în cazul utilizării ambelor substanţe.

Tabel de certitudine a

Cînd mărimea indicelui mai mic din cei 2 indici este de ...%	Cînd numărul de observaţii care stau							
	20	30	40	50	60	70	80	90
1	19	14	11	9,0	7,9	6,9	6,2	5,7
2	20	15	12	10	8,9	7,9	7,2	6,7
3	21	16	13	11	9,8	8,8	8,0	7,4
4	22	17	14	12	11	9,5	8,7	8,0
5	23	18	14	13	11	10	9,3	8,6
6	24	18	15	13	12	11	9,8	9,1
7	24	19	16	14	12	11	10	9,6
8	25	19	16	14	13	12	11	10
9	25	20	17	15	13	12	11	10
10	26	20	17	15	14	12	11	11
15	28	22	19	17	15	14	13	12
20	29	24	20	18	16	15	14	13
25	30	25	21	19	17	16	15	14
30	31	25	22	19	18	16	15	14
35	31	25	22	20	18	17	16	15
40	31	25	22	20	18	17	16	15
45	30	25	22	20	18	17	16	15
50	30	25	22	19	18	17	16	15
55	29	24	21	19	17	16	15	14
60	27	23	19	18	17	16	15	14
65	26	22	19	17	16	15	14	13
70	24	20	18	16	15	14	13	13
75	22	19	17	15	14	13	12	12
80	18	16	15	13	13	12	11	10
85	—	14	13	12	11	10	9,5	9,0
90	—	—	10	9,0	8,5	8,0	7,6	7,2
91	—	—	9,1	8,4	7,9	7,5	7,1	6,8
92	—	—	8,4	7,8	7,4	7,0	6,7	6,3
93	—	—	—	7,0	6,8	6,4	6,1	5,8
94	—	—	—	—	—	5,8	5,5	5,2
95	—	—	—	—	—	—	4,9	4,7
96	—	—	—	—	—	—	—	4,0
97	—	—	—	—	—	—	—	—
98	—	—	—	—	—	—	—	—
99	—	—	—	—	—	—	—	—

Numărul mare de animale de laborator, atît în lotul experimental, cît şi în lotul martor ne permite să aplicăm legea numerelor mari în calcularea valorilor medii, cît şi a dispersiunii de la medie.

Calculăm, în prealabil, indicii de mortalitate. În lotul experimental indicele de mortalitate va fi 55,0%, iar la lotul martor 20,0%.

$$\frac{33 \cdot 100}{60} = 55{,}0\% \qquad \frac{6 \cdot 100}{30} = 20{,}0\%$$

Aplicăm formula cunoscută cu ocazia stabilirii dispersiunii de la medie în cazul frecvenţelor relative :

$$\sigma_1 = \pm \sqrt{\frac{P_1 \cdot Q_1}{N_1}} = \pm \sqrt{\frac{P_1 \cdot (100 - P_1)}{N_1}} = \pm \sqrt{\frac{55.45}{60}} = \pm 6{,}42$$

$$\sigma_2 = \pm \sqrt{\frac{P_2 \cdot Q_2}{N_2}} = \pm \sqrt{\frac{P_2 \cdot (100 - P_2)}{N_2}} = \pm \sqrt{\frac{20.80}{30}} = \pm 7{,}30$$

Tabelul XII

diferenţei dintre doi indici

la baza fiecărui indice este de :

100	200	300	400	500	600	700	800	900	1 000
5,3	3,2	2,4	2,0	1,7	1,5	1,4	1,3	1,0	0,9
6,3	3,9	3,0	2,5	2,2	1,9	1,8	1,7	1,6	1,5
6,9	4,5	3,4	2,9	2,6	2,3	2,1	2,0	1,8	1,7
7,5	4,9	3,8	3,2	2,9	2,6	2,4	2,2	2,1	1,9
8,1	5,3	4,2	3,5	3,1	2,8	2,6	2,4	2,3	2,1
8,5	5,6	4,5	3,8	3,4	3,1	2,8	2,6	2,5	2,3
9,0	6,0	4,7	4,0	3,6	3,3	3,0	2,8	2,7	2,4
9,4	6,3	5,0	4,3	3,8	3,4	3,2	2,9	2,8	2,5
9,7	6,5	5,2	4,5	4,0	3,6	3,3	3,1	2,9	2,7
10	6,8	5,4	4,6	4,1	3,7	3,4	3,2	3,0	2,8
11	7,8	6,3	5,4	4,8	4,3	4,0	3,7	3,5	3,3
12	8,6	6,9	5,9	5,3	4,8	4,4	4,1	3,9	3,7
13	9,1	7,4	6,3	5,7	5,2	4,8	4,4	4,2	4,0
14	9,5	7,7	6,6	5,9	5,4	5,0	4,7	4,2	4,1
14	9,8	7,9	6,9	6,1	5,6	5,2	4,8	4,5	4,3
14	9,9	8,1	7,0	6,4	5,7	5,3	4,9	4,6	4,4
14	10	8,1	7,1	6,3	5,7	5,3	5,0	4,7	4,5
14	9,9	8,1	7,0	6,3	5,7	5,3	5,0	4,7	4,5
14	9,7	8,0	6,9	6,2	5,6	5,2	4,9	4,6	4,4
13	9,5	7,8	6,8	6,2	5,5	5,1	4,8	4,5	4,4
13	9,2	7,6	6,5	5,9	5,4	5,0	4,7	4,4	4,2
12	8,7	7,1	6,2	5,6	5,2	4,7	4,5	4,2	4,2
11	8,1	6,7	5,8	5,3	4,9	4,4	4,2	4,0	3,8
10	7,3	6,1	5,3	4,8	4,4	4,1	3,8	3,6	3,4
7,9	6,4	5,3	4,7	4,2	3,8	3,6	3,4	3,2	3,0
6,9	5,2	4,4	3,8	3,5	3,2	3,0	2,8	2,6	2,5
6,5	4,9	4,1	3,6	3,3	3,0	2,8	2,6	2,5	2,4
6,1	4,6	3,9	3,4	3,1	2,9	2,7	2,5	2,3	2,0
5,6	4,2	3,6	3,2	2,9	2,7	2,5	2,4	2,2	2,1
5,1	3,8	3,3	3,0	2,6	2,5	2,3	2,2	2,1	1,9
4,5	3,5	3,0	2,6	2,4	2,2	2,1	1,9	1,8	1,8
3,9	3,0	2,6	2,3	2,1	2,0	1,9	1,7	1,6	1,6
—	2,7	2,2	2,0	1,8	1,7	1,6	1,5	1,5	1,3
—	2,0	1,8	1,5	1,4	1,3	1,2	1,2	1,1	1,0
—	—	—	1,0	0,9	0,8	0,8	0,7	0,6	0,5

Calculăm, în continuare, eroarea medie a mediei aritmetice :

$$e\,M_1 = \pm \frac{\sigma_1}{\sqrt{N_1}} = \pm \frac{6{,}42}{\sqrt{60}} = \pm 0{,}83$$

$$e\,M_2 = \pm \frac{\sigma_2}{\sqrt{N_2}} = \pm \frac{7{,}30}{\sqrt{30}} = \pm 1{,}33$$

Aplicînd triplul erorii medii, ne putem da seama că mortalitatea la lotul experimental se cuprinde între 55,0% $\pm$ 3.eM_1 adică 55,0% $\pm$3.0,83 = 52,51% şi 57,49%.

Mortalitatea la lotul martor se cuprinde între 20,0% $\pm$ 3.eM_2, adică 20,0% $\pm$ 3.1,33 = 16,01% şi 23,99%.

Pentru certitudinea diferenţei, adică pentru a ne da seama dacă diferenţa dintre indicii de letalitate este sau nu semnificativă, deci dacă în cazul nostru acţiunea celor două substanţe A şi B este sau nu toxică, calculăm indicele de eroare a diferenţei.

Diferenţa dintre cei doi indici de mortalitate este de 35,0% (55,0% —20,0% = 35,0%).

Eroarea diferenţei se calculează după formula cunoscută :

$$e\,D = \pm \sqrt{e\,M_1^2 + e\;M_2^2} = \pm \sqrt{0{,}83 + 1{,}33} = \pm 1{,}47$$

Triplul erorii diferenţei va fi de 4,41, indice mult inferior diferenţei dintre cele două medii.

$$35\% > 4{,}41\%$$

Deci avem certitudinea că diferenţa între mortalitatea celor două loturi nu este întîmplătoare şi că acţiunea substanţei A este toxică, iar cea a substanţei B antitoxică.

Există o serie de tabele prin care se poate deduce mult mai uşor certitudinea diferenţei, fără calculele expuse anterior.

Amintim astfel tabelele apărute în editura Academiei de ştiinţe medicale din U.R.S.S.

Să presupunem, în cazul nostru, că lotul martor ar fi fost egal ca număr cu lotul experimental, şi anume 70, indicii de mortalitate rămînînd aceiaşi pe care i-am văzut : 55,0% şi 20,0%.

Să vedem în ce măsură putem considera diferenţele dintre cele două medii de letalitate ca esenţiale, deci să confirmăm eficienţa celor două substanţe A şi B.

Căutînd în tabelul XII, în coloana I, pe verticală, cifra ce corespunde indicelui mai mic de mortalitate, în cazul nostru 20,0%, iar pe orizontală coloana ce corespunde frecvenţei, adică numărul de animale ce alcătuieşte fiecare dintre loturi, în cazul nostru 70. La intersecţia rîndului orizontal şi a coloanei verticale citim cifra 15.

Aceasta înseamnă că dacă diferența celor doi indici de letalitate este de peste 15,0%, ea este semnificativă. În cazul nostru diferența dintre cei doi indici de mortalitate fiind de 35,0%, putem confirma certitudinea diferenței, deci și certitudinea acțiunii antagonice a celor două substanțe : A și B.

Această diferență semnificativă a indicilor este calculată în tabel la 2 σ, deci cu o probabilitate de 95,45%. În caz că, pentru o mai mare exactitate, utilizăm 3 σ (probabilitatea de 99,73%), vom înmulți cifra obținută din tabel cu 1,5. În cazul nostru, înmulțind 1,5 cu 15, vom obține 22,5, indice mai mic decît diferența dintre cele două mortalități (35,0%) la lotul experimental și lotul martor. Deci, și în acest caz, certitudinea diferenței este demonstrată și statistic.

Tot din tabel ne putem da seama că, pe lîngă o diferență atît de mare dintre mortalitatea lotului experimental și lotul martor, chiar și un eșantion de cîte 30 de animale pentru fiecare lot ar fi fost semnificativ. Intr-adevăr la intersecția rîndului 20, corespunzător indicelui de mortalitate de 20,0% a lotului martor cu coloana 30 (număr de animale din fiecare lot), găsim valoarea 24 (utilizînd 2 σ), valoare mult inferioară diferenței dintre mortalitatea celor două loturi, care este 35,0%. În cazul utilizării a 3 σ, deci pentru o certitudine de 99,73%, vom găsi o valoare de 36,0% ($\pm 1{,}5.24 = 36{,}0\%$), valoare doar cu puțin mai mare decît diferența dintre cei doi indici de mortalitate.

4. CALCULUL CERTITUDINII DIFERENȚEI DINTRE DOI INDICI PRIN METODA χ^2 (CHI PĂTRAT)

Prin metoda χ^2 putem, de asemenea, calcula cu ușurință certitudinea diferenței dintre două medii. În acest caz utilizăm tabelele cvadruple, pe care le-am cunoscut anterior (tabelul XIII).

Tabelul XIII

	Au decedat	N-au decedat	Total
Lotul experimental	$33_{(a_1)}$	$27_{(b_1)}$	60 $(a_1 + b_1 = n_1)$
Lotul martor	$6_{(a_2)}$	$24_{(b_2)}$	30 $(a_2 + b_2 = n_2)$
Total	39 $(a_1 + a_2 = n_3)$	51 $(b_1 + b_2 = n_4)$	90 $(n_1 + n_2 = n_3 + n_4 = N)$

$$\chi^2 = \frac{(a_1 \cdot b_2 - b_1 \cdot a_2)^2}{n_1 \cdot n_2 \cdot n_3 \cdot n_4} \cdot (N - 1)$$

$$\chi^2 = \frac{(33 \cdot 24 - 27 \cdot 6)^2}{60 \cdot 30 \cdot 39 \cdot 51} \cdot (90 - 1) = 9{,}87$$

χ^2 fiind peste valoarea 9 certitudinea diferenței este demonstrată cu o probabilitate de peste 99,73%.

Din cele arătate anterior ne putem da seama că nu este necesar să utilizăm un număr atît de important de animale de laborator.

În general, cu cît ne aşteptăm — pe baza experimentelor anterioare — la diferențe mai mari între acțiunea unei substanțe asupra lotului experimental față de lotul martor, cu atît mărimea lotului poate fi mai mică.

În caz că lotul este mai mic de 30 nu mai putem aplica calculul bazat pe distribuția normală. Vom aplica însă în acest caz metoda χ^2, care la eşantioane de 5 sau mai bine de 10 animale de laborator poate să ne demonstreze certitudinea diferenței.

Să luăm un exemplu similar cu cel anterior, în care să difere doar mărimea eşantionului, lotul experimental fiind compus din 24 de animale de laborator, iar lotul martor din 12 animale.

Administrînd substanța A, respectiv la lotul martor substanța A şi B, am obține, spre exemplu, 16 decese la lotul experimental şi 3 decese la lotul martor. Mortalitatea ar fi, deci, la lotul experimental de 66,7%, iar la lotul martor de 25,0%.

Diferența de mortalitate este evidentă, totuşi să aplicăm şi aici calculul certitudinii diferenței dintre cei doi indici.

Vom aplica metoda χ^2, deoarece numărul eşantioanelor este sub 30 (tabelul XIV).

Tabelul XIV

	Au decedat	N-au decedat	Total
Lotul experimental	16	8	24
Lotul martor	3	9	12
Total	19	17	36

$$\chi^2 = \frac{(16 \cdot 9 - 8 \cdot 3)^2}{24 \cdot 12 \cdot 19 \cdot 17} \cdot (36 - 1) = 5{,}42.$$

Din cele arătate în capitolul I ştim că atunci cînd χ^2 este mai mare de 9, certitudinea diferenței este demonstrată cu o probabilitate de 99,7%, cînd depăşeşte 6,64 probabilitatea este de 99,0%, iar la un χ^2 egal cu 3,84 certitudinea diferenței este demonstrată cu o probabilitate de 95,0%.

În cazul nostru cu o probabilitate de aproape 99,0% am demonstrat că diferența între cei doi indici de mortalitate este certă, prin acest fapt demonstrînd şi certitudinea eficacității celor două substanțe A şi B.

5. CALCULUL DISPERSIUNII CU AJUTORUL TABELELOR

Pe baza tabelelor elaborate de Academia de ştiinţe medicale din U.R.S.S. se poate calcula cu uşurinţă dublul deviaţiei tip a mediilor în cazul în care eşantionul nu este mai mic de 20. De altfel se ştie că este mai bine ca eşantionul minim să fie de cel puţin 30.

Din tabelul XV, calculăm dublul deviaţiei-tip, cunoscînd valoarea medie şi numărul eşantionului.

Luăm exemplul privitor la lotul experimental compus din 60 de animale de laborator şi cu un indice de mortalitate de 55,0%, iar lotul martor alcătuit din 30 de animale, cu o mortalitate de 20,0%.

Căutăm în tabel cifra de intersecţie a rîndului care corespunde indicelui de mortalitate la lotul experimental 55%, cu coloana ce corespunde frecvenţei de 60 (numărul de animale) şi vom găsi dublul deviaţiei tip egal cu ± 13, apropiat de valoarea $\pm 6{,}42$ pe care am aflat-o prin calculul anterior şi care dublată va fi $\pm 12{,}84$.

În cazul lotului martor, la intersecţia rîndului care corespunde indicelui de 20,0% (indice de mortalitate) cu coloana 30 (numărul de animale martor), vom găsi, în acelaşi tabel, valoarea ± 15, corespunzătoare dublului deviaţiei tip, care, pe baza calculului anterior, a fost de $\pm 7{,}30$, care dublat va fi 14,6.

Deci prin această metodă, putem afla cu multă operativitate dublul deviaţiei standard.

Dacă pentru o mare precizie nu ne mulţumim cu probabilitatea de 95,45%, pe care ne-o dă dublul deviaţiei standard, putem lua întreitul ei, în care caz vom avea, după cum s-a mai arătat, o probabilitate mult mai mare de 99,73%.

În această situaţie, însă, pentru a putea utiliza tabelul XV se adaugă valorii găsite prin metoda anterioară încă o jumătate din valoarea respectivă. Deci în loc de ± 12 vom avea ± 18, iar în al doilea caz în loc de ± 15 vom avea $\pm 22{,}5$; triplul valorilor calculate de noi anterior ar fi $\pm$ 3.6,42 = $\pm$ 19,26, iar în cazul al doilea $\pm 3.7{,}3 = \pm 21{,}9$, deci valori apropiate de cele găsite în tabel.

6. CALCULUL DISPERSIUNII ÎN CAZUL EVENIMENTELOR RARE

Sînt cazuri cînd un experiment oarecare se soldează printr-un efect slab, cînd deci probabilitatea ca un efect oarecare să se producă este rară. În aceste cazuri, vom putea calcula probabilitatea de apariţie a fenomenului aşteptat, chiar şi atunci cînd în cazul eşantionului utilizat el nu a apărut încă, adică a fost zero.

Se utilizează, în aceste situaţii, distribuţia Poisson în locul distribuţiei normale, distribuţie cunoscută şi sub numele de „legea numerelor mici".

În cazul evenimentelor rare se utilizează formula şi graficul din fig. 7.

În grafic putem afla cu uşurinţă diferitele distribuţii, cînd cunoaştem valoarea care corespunde probabilităţii de apariţie a fenomenului.

Calculul

Cînd mărimea indicelui este de ..%	Cînd numărul de							
	20	30	40	50	60	70	80	90
1	—*	—	—	—	—	—	—	—
2	—	—	—	—	—	—	—	—
3	—	—	—	—	—	—	—	—
4	—	—	—	—	—	—	—	—
5	—	—	—	—	—	—	—	—
6	—	—	—	—	—	—	4,9	4,6
7	—	—	—	—	—	5,7	5,3	5,0
8	—	—	—	—	6,6	6,2	5,7	5,4
9	—	—	—	7,8	7,1	6,6	6,1	5,7
10	—	—	—	8,2	7,4	6,9	6,4	6,0
15	—	—	9,6	8,6	7,8	7,2	6,8	6,3
20	18	13	11	10	9,3	8,6	8,0	7,5
25	20	15	13	11	10	9,6	9,0	8,4
30	21	16	14	12	11	10	9,7	9,1
35	22	17	15	13	12	11	10	9,7
40	23	18	15	14	12	12	11	10
45	23	18	16	14	13	12	11	10
50	23	18	16	14	13	12	11	10
55	23	19	16	14	13	12	11	10
60	23	18	16	14	13	12	11	10
65	23	18	16	14	13	12	11	10
70	22	18	15	14	12	12	11	10
75	21	17	15	13	12	11	10	9,7
80	20	16	14	12	11	10	9,7	9,1
85	18	15	13	11	10	9,6	9,0	8,4
90	16	13	11	10	9,3	8,6	8,0	7,5
91	14	11	9,6	8,6	7,8	7,2	6,8	6,3
92	13	11	9,2	8,2	7,4	6,9	6,4	6,0
93	13	10	8,7	7,8	7,1	6,6	6,1	5,7
94	12	9,5	8,2	7,3	6,6	6,2	5,7	5,4
95	11	8,8	7,6	6,8	6,2	5,7	5,3	5,0
96	10	8,1	7,0	6,2	5,7	5,3	4,9	4,6
97	9,0	7,3	6,3	5,6	5,1	4,7	4,4	4,1
98	7,9	6,3	5,5	4,4	4,5	4,1	3,9	3,6
99	6,4	5,2	4,5	4,0	3,6	3,4	3,2	3,0
	4,5	3,7	3,2	2,8	2,6	2,4	2,2	2,1

* Înseamnă că limita erorii depăşeşte indicele însăşi.

Dacă, spre exemplu, unui lot de 10 animale de laborator i se administrează o substanţă care provoacă anumite schimbări în organismul lor, putem să calculăm care este procentul maxim cu care s-ar fi produs aceste schimbări, în cazul în care ar fi utilizat un lot mai mare.

Citim, în graficul din fig. 7, în dreptul valorii lui n egal în cazul nostru cu 10 (numărul de animale din lot), ce mărime îi corespunde lui $p_{o,n}$ şi vom găsi 4,4.

Aplicînd formula lui $m = \frac{p_{o,n}}{n}$, vom găsi $\frac{4,4}{10} = 0,44$, iar în cazul calculului procentual al lui m : $m\% = 100.0,44 = 44\%$.

Deci se poate afirma cu o probabilitate de 99,73% că în cazul observării unui lot mult mai mare, posibilitatea de a se produce schim-

Tabelul XV

dispersiunii (2σ)

observaţii este :

100	200	300	400	500	600	700	800	900	1000
—	—	—	1,0	0,9	0,8	0,8	0,7	0,7	0,6
—	2,0	1,6	1,4	1,3	1,1	1,1	1,0	0,9	0,9
—	2,4	2,0	1,7	1,5	1,4	1,3	1,2	1,1	1,1
3,9	2,8	2,3	2,0	1,8	1,6	1,5	1,4	1,3	1,2
4,0	3,1	2,5	2,2	2,0	1,8	1,7	1,5	1,5	1,4
4,8	3,4	2,7	2,4	2,1	1,9	1,9	1,7	1,6	1,5
5,1	3,6	2,9	2,6	2,3	2,1	1,9	1,8	1,7	1,6
5,5	3,9	3,1	2,7	2,4	2,2	2,1	1,9	1,8	1,7
5,8	4,1	3,3	2,9	2,6	2,3	2,2	2,0	1,9	1,8
6,0	4,3	3,5	3,0	2,7	2,5	2,3	2,1	2,0	1,9
7,2	5,1	4,1	3,6	3,2	2,9	2,7	2,5	2,4	2,3
8,0	5,7	4,6	4,0	3,6	3,3	3,0	2,8	2,7	2,5
8,7	6,2	5,0	4,3	3,9	3,5	3,3	3,1	2,9	2,7
9,2	6,5	5,3	4,6	4,1	3,7	3,5	3,2	3,1	2,9
9,6	6,8	5,5	4,8	4,3	3,9	3,6	3,4	3,2	3,1
9,9	7,0	5,6	4,9	4,4	4,0	3,7	3,5	3,3	3,1
10	7,1	5,7	5,0	4,5	4,1	3,8	3,5	3,3	3,2
10	7,1	5,8	5,0	4,5	4,1	3,8	3,5	3,3	3,2
9,9	7,0	5,6	4,9	4,4	4,0	3,7	4,5	3,3	3,1
9,6	6,8	5,5	4,8	4,3	3,9	3,6	3,4	3,2	3,0
9,2	6,5	5,3	4,6	4,1	3,7	3,5	3,2	3,1	2,9
8,7	6,2	5,0	4,3	3,9	3,5	3,3	3,1	2,9	2,7
8,0	5,7	4,6	4,0	3,6	3,3	3,0	2,8	2,7	2,5
7,2	5,1	4,1	3,6	3,2	2,9	2,7	2,5	2,4	2,3
6,0	4,3	3,5	3,0	2,7	2,5	2,3	2,1	2,0	1,9
5,8	4,1	3,3	2,9	2,6	2,3	2,2	2,0	1,9	1,8
5,5	3,9	3,1	2,7	2,4	2,2	2,1	1,9	1,8	1,7
5,1	3,6	2,9	2,6	2,3	2,1	1,9	1,8	1,7	1,6
4,8	3,4	2,7	2,4	2,1	1,9	1,8	1,7	1,6	1,5
4,4	3,1	2,5	2,2	2,0	1,8	1,7	1,5	1,5	1,4
3,9	2,8	2,3	2,0	1,8	1,6	1,5	1,4	1,3	1,2
3,4	2,4	2,0	1,7	1,5	1,4	1,3	1,2	1,1	1,1
2,8	2,0	1,6	1,4	1,3	1,1	1,1	1,0	0,9	0,9
2,0	1,4	1,1	1,0	0,9	0,8	0,8	0,7	0,7	0,6

bările cu ocazia administrării substanţei respective, n-ar fi depăşit proporţia de 44,0% din numărul animalelor de laborator utilizate.

7. CALCULUL EŞANTIONULUI REPREZENTATIV

Putem calcula cu multă uşurinţă eşantionul necesar, numărul de animale de laborator dintr-un lot experimental pe baza tabelului XVI.

Să luăm un lot de 50 de animale de laborator care vaccinate cu un vaccin anumit (X) au devenit imune într-o proporţie de 85% faţă de un lot identic ca mărime, vaccinat cu vaccin Y, la care imunitatea s-a ivit într-o proporţie de 65%.

Tabelul XV

dispersiunii (2σ)

observaţii este :

100	200	300	400	500	600	700	800	900	1000
—	—	—	1,0	0,9	0,8	0,8	0,7	0,7	0,6
—	2,0	1,6	1,4	1,3	1,1	1,1	1,0	0,9	0,9
—	2,4	2,0	1,7	1,5	1,4	1,3	1,2	1,1	1,1
3,9	2,8	2,3	2,0	1,8	1,6	1,5	1,4	1,3	1,2
4,0	3,1	2,5	2,2	2,0	1,8	1,7	1,5	1,5	1,4
4,8	3,4	2,7	2,4	2,1	1,9	1,9	1,7	1,6	1,5
5,1	3,6	2,9	2,6	2,3	2,1	1,9	1,8	1,7	1,6
5,5	3,9	3,1	2,7	2,4	2,2	2,1	1,9	1,8	1,7
5,8	4,1	3,3	2,9	2,6	2,3	2,2	2,0	1,9	1,8
6,0	4,3	3,5	3,0	2,7	2,5	2,3	2,1	2,0	1,9
7,2	5,1	4,1	3,6	3,2	2,9	2,7	2,5	2,4	2,3
8,0	5,7	4,6	4,0	3,6	3,3	3,0	2,8	2,7	2,5
8,7	6,2	5,0	4,3	3,9	3,5	3,3	3,1	2,9	2,7
9,2	6,5	5,3	4,6	4,1	3,7	3,5	3,2	3,1	2,9
9,6	6,8	5,5	4,8	4,3	3,9	3,6	3,4	3,2	3,1
9,9	7,0	5,6	4,9	4,4	4,0	3,7	3,5	3,3	3,1
10	7,1	5,7	5,0	4,5	4,1	3,8	3,5	3,3	3,2
10	7,1	5,8	5,0	4,5	4,1	3,8	3,5	3,3	3,2
10	7,1	5,7	5,0	4,5	4,1	3,8	3,5	3,3	3,2
9,9	7,0	5,6	4,9	4,4	4,0	3,7	4,5	3,3	3,1
9,6	6,8	5,5	4,8	4,3	3,9	3,6	3,4	3,2	3,0
9,2	6,5	5,3	4,6	4,1	3,7	3,5	3,2	3,1	2,9
8,7	6,2	5,0	4,3	3,9	3,5	3,3	3,1	2,9	2,7
8,0	5,7	4,6	4,0	3,6	3,3	3,0	2,8	2,7	2,5
7,2	5,1	4,1	3,6	3,2	2,9	2,7	2,5	2,4	2,3
6,0	4,3	3,5	3,0	2,7	2,5	2,3	2,1	2,0	1,9
5,8	4,1	3,3	2,9	2,6	2,3	2,2	2,0	1,9	1,8
5,5	3,9	3,1	2,7	2,4	2,2	2,1	1,9	1,8	1,7
5,1	3,6	2,9	2,6	2,3	2,1	1,9	1,8	1,7	1,6
4,8	3,4	2,7	2,4	2,1	1,9	1,8	1,7	1,6	1,5
4,4	3,1	2,5	2,2	2,0	1,8	1,7	1,5	1,5	1,4
3,9	2,8	2,3	2,0	1,8	1,6	1,5	1,4	1,3	1,2
3,4	2,4	2,0	1,7	1,5	1,4	1,3	1,2	1,1	1,1
2,8	2,0	1,6	1,4	1,3	1,1	1,1	1,0	0,9	0,9
2,0	1,4	1,1	1,0	0,9	0,8	0,8	0,7	0,7	0,6

bările cu ocazia administrării substanţei respective, n-ar fi depăşit proporţia de 44,0% din numărul animalelor de laborator utilizate.

7. CALCULUL EŞANTIONULUI REPREZENTATIV

Putem calcula cu multă uşurinţă eşantionul necesar, numărul de animale de laborator dintr-un lot experimental pe baza tabelului XVI.

Să luăm un lot de 50 de animale de laborator care vaccinate cu un vaccin anumit (X) au devenit imune într-o proporţie de 85% faţă de un lot identic ca mărime, vaccinat cu vaccin Y, la care imunitatea s-a ivit într-o proporţie de 65%.

Să vedem dacă numărul de animale utilizate sînt suficiente pentru a constata eficiența mai mare a vaccinului X față de vaccinul Y sau dacă această diferență este doar întîmplătoare.

Pentru a avea în cazul nostru această certitudine trebuie să considerăm o eroare medie de cel mult 10% pentru că în acest caz dublul erorii medii care este de 2.10% = 20% să nu depășească diferența dintre cei doi indici de imunizare, căci altfel diferența nu mai este certă (85%—65% = 20%, care este egal cu dublul erorii medii 2.10%).

Acestea fiind fixate căutăm în tabelul XVI valoarea ce corespunde intersecției dintre rîndul orizontal 85 (proporția de imunizare cu vaccinul X presupus mai eficace) și coloana verticală 10 (eroarea medie admisă de noi ca maximă).

Tabelul XVI

Numărul de observații necesare pentru ca eroarea să nu depășească limita dată de 19 cazuri din 20

Cînd mărimea indicelui este de ... %	Limita erorii relative în %					
	1	2	3	4	5	10
1	400	100	45	26	17	5
2	800	200	90	50	32	9
3	1 200	300	130	74	48	13
4	1 500	400	170	100	62	16
5	1 900	500	210	120	77	20
6	2 300	600	250	140	90	24
7	2 600	650	290	160	110	27
8	3 000	740	330	190	120	30
9	3 300	800	370	210	130	34
10	3 600	900	400	230	150	37
15	5 100	1 300	570	320	210	52
20	6 400	1 600	710	400	260	65
25	7 500	1 900	830	470	300	76
30	8 400	2 100	930	530	340	85
35	9 100	2 300	1 010	570	370	92
40	9 600	2 400	1 070	600	390	97
45	9 900	2 500	1 100	620	400	100
50	10 000	2 500	1 100	630	400	100
55	9 900	2 500	1 100	620	400	100
60	9 600	2 400	1 070	600	390	97
65	9 100	2 300	1 010	570	370	92
70	8 400	2 100	930	530	340	85
75	7 500	1 900	830	470	300	76
80	6 400	1 600	710	400	260	65
85	5 100	1 300	570	320	210	52
90	3 600	900	400	230	150	37
91	3 300	800	370	210	130	34
92	3 000	740	330	190	120	30
93	2 600	650	290	160	110	27
94	2 300	600	250	140	90	24
95	1 900	500	210	120	77	20
96	1 500	400	170	100	62	16
97	1 200	300	130	74	48	13
98	800	200	90	50	32	9
99	400	100	45	26	17	5

Vom găsi valoarea 52, ceea ce înseamnă că este necesar un eșantion de 52 de animale de laborator pentru a demonstra superioritatea vaccinului X față de vaccinul Y.

În cazul nostru, avînd un număr aproape identic cu eșantionul reprezentativ (50 de animale) putem afirma cu o probabilitate de 95,45% superioritatea vaccinului X.

Pentru o mai mare certitudine (99,73% ce ar corespunde triplului erorii medii admise) trebuie să creștem aproape la dublu numărul eșantionului, să avem adică 100 de animale de laborator.

Specificăm, însă, că diferența de imunizare dintre cele 2 vaccinuri este atît de mare (85%—65%), încît putem fi mulțumiți cu eșantionul luat de noi (cîte 50 de animale), corespunzînd unei probabilități de 95,45%.

CAPITOLUL al III-lea

METODE MATEMATICE ÎN CERCETĂRILE CLINICE

1. NOȚIUNI GENERALE

Ținînd seama de faptul că observația clinică este metoda de bază în cercetarea sănătății sau îmbolnăvirii organismului (pe lîngă metoda de laborator), rezultatele obținute se caracterizează prin marea lor variabilitate.

Organismul reacționează în mod diferit la impulsurile venite din mediul extern și intern, impulsuri care, la rîndul lor, sînt și ele mult variate ca intensitate.

Nu putem vorbi, de exemplu, de o entitate nosologică absolut uniformă la toți bolnavii de o anumită boală, după cum nu putem vorbi nici de elementele fiziologice care ar caracteriza, în mod precis, pe omul „sănătos"; există o scară întreagă de valori diferite, care dau aspectul de individualitate a cazurilor clinice observate.

Cu toată variabilitatea (dispersiunea) cazurilor individuale, există o serie de însușiri comune, care fac, totuși, posibilă clasificarea cazurilor similare într-o grupă mai mult sau mai puțin omogenă, mergînd pe aspectele cele mai importante ce caracterizează omul sănătos sau omul bolnav.

Această mare variabilitate a cazurilor clinice face să abunde în literatură numeroși termeni vagi, în ceea ce privește simptomatologia multor boli sau evoluția acestora, ceea ce îngreuiază, fie punerea diagnosticului sau stabilirea prognosticului, fie delimitarea precisă a bolii și sănătății.

În astfel de cazuri (de mare variabilitate a fenomenelor) tocmai statistica matematică este aceea care permite să se efectueze, în mod științific, grupările omogene, pe baza cărora să se poată trage concluziile necesare pentru stabilirea diagnosticului, a comportamentului terapeutic și a prognosticului. Concluziile trase numai pe baza observației clinice, neînsoțite de gruparea statistică și fără a se calcula indicii medii, de dispersiune, de eroare medie, mărimea eșantionului reprezentativ etc. riscă, de multe ori, să fie eronate.

Există, desigur, mari greutăți în încadrarea în grupe omogene a observațiilor clinice, ținînd seamă nu numai de variabilitatea indivizilor, ci și de bolile asociate, complicațiile bolilor respective (toate acestea în funcție și de modul de reactivitate a organismului).

Dacă în medicina experimentală cercetătorul poate reproduce la voința sa condiții aproape similare în repetarea experimentelor în laborator, clinicianul nu poate decît rareori să-și creeze condiții identice în cadrul observației pe mai multe persoane ce se încadrează într-o anumită boală. El trebuie, în general, să facă observația pe cazurile ce se ivesc în afara voinței sale și chiar dacă există un număr mai mare de cazuri de aceeași afecțiune, condițiile, formele de boală nu sînt în general identice. Din aceste considerente și rezultatele observației vor suferi ca exactitate, atunci cînd cercetătorul caută să clasifice cazurile observate.

Această situație îngreuiază și comparabilitatea cazurilor individuale, adică compararea acestora din punct de vedere al debutului, evoluției, decesului, vindecării, agravării, ameliorării etc. cu o „boală-tip", în vederea punerii diagnosticului, a instituirii tratamentului și a stabilirii unui prognostic.

Dar dacă această comparabilitate a cazurilor individuale cu un „etalon" nu este posibilă în general, ea nici nu este necesară, deoarece, cu toată variabilitatea amintită, există întotdeauna caracteristici esențiale, comune categoriilor de boli (etiologie, modificări fiziopatologice și anatomice etc.), care permit formarea unor grupe omogene.

Din considerentele anterior amintite, în cazul observației clinice nu se utilizează loturi martore, deși această utilizare este posibilă în multe cazuri. Astfel utilizarea unor preparate medicamentoase pe un anumit lot de bolnavi și a altor preparate pe un alt lot ce servește drept martor sau asocierea diferitelor preparate medicamentoase, sau a diferitelor proceduri terapeutice la anumite categorii de bolnavi, nu numai că este posibilă, dar este și necesară. În acest caz, însă, am asociat observației clinice și metoda experimentală.

Asocierea observației clinice cu experimentul și cu investigațiile de laborator îi permite clinicianului nu numai stabilirea diagnosticului și prognosticului, precum și instituirea tratamentului, dar și clasificarea cazurilor observate, generalizarea acestor observații în vederea progresului științei medicale. Constatările logice, pe calea observației și a experimentului, pe calea inductivă, permit cercetătorului să ajungă la cunoașterea legăturilor de cauzalitate în ceea ce privește ivirea bolii, trecerea de la sănătate la boală și invers.

Corelarea observației clinice cu investigațiile de laborator și cu experimentul, este, deci, obligatorie în cercetările clinice. Cu atît mai mult, această corelare necesită utilizarea statisticii, care devine un mijloc de verificare a exactității rezultatelor obținute din cercetarea clinică.

Cu ajutorul metodelor matematice putem să deducem valorile medii și gradul de dispersiune de la această valoare medie, limitele posibile între care cazurile individuale dintr-o categorie anumită de boală pot diferi față de această valoare medie, fără ca să fie considerate cazuri deosebite. De asemenea, metodele matematice ne pot evidenția posibilitatea unei mortalități mărite chiar și în cazul în care la un număr mic de observații nu s-au ivit decese și, invers, cînd la un număr mic de cazuri observate acestea toate s-au soldat prin deces, statistica matematică ne poate arăta indicele de mortalitate care s-ar fi obținut pe un număr suficient de mare de cazuri și totodată limita superioară a acestui indice.

În cazul comparării a doi indici de mortalitate sau în cazul examenelor repetate de laborator, metodele matematice pot scoate în evidență

certitudinea diferenței diferitelor rezultate sau, din contră, pot arăta că diversele rezultate obținute sînt doar întîmplător diferite, din cauza unei variabilități mai mari a cazurilor observate, a lipsei omogenității sau din cauza unei grupări necorespunzătoare.

Cele mai multe lipsuri în ceea ce privește observația clinică se leagă de numărul scăzut al cazurilor observate și din această cauză generalizările care se fac nu corespund în bună parte. Cîte cazuri trebuie observate pentru a putea ajunge la concluzii general-valabile? Metodele matematice ne pot da un răspuns concret prin indicele mărimii reprezentative a eșantionului.

Deși numărul mare de observații ne duce la concluzii general-valabile, conform legii numerelor mari, totuși, de multe ori, nu putem dispune, după necesități, de un număr suficient de cazuri ce se pot lua în observație, cum ar fi, de exemplu, în bolile rare. Înseamnă că nu se pot generaliza observațiile efectuate la un număr mai mic de cazuri?

Statistica matematică ne ajută să dăm un răspuns pozitiv și în această problemă: cu cît cazurile observate sînt mai omogene, cu atît numărul de observații poate fi mai redus, bineînțeles nu însă sub o anumită limită.

Metoda χ^2 ne ajută să tragem concluzii chiar dintr-un număr de 10 cazuri observate, cu condiția ca ele să fie omogene. Distribuția Poisson, așa-numita „lege a numerelor mici", ne ajută, de asemenea, în rezolvarea cercetărilor pe cazuri rare.

Omogenitatea cazurilor cercetate, nu poate însă suplini întotdeauna un număr suficient de mare de cazuri cercetate, deoarece riscăm ca eșantionul să nu mai fie reprezentativ și invers, crescînd mult numărul de cazuri observate riscăm ca aceasta să meargă în dauna omogenității. *Ca regulă generală se poate stabili că în cazurile rare dar omogene putem să tragem concluzii general-valabile și pe baza unui număr scăzut de observații (cel puțin 10 cazuri din boala respectivă), dar atunci cînd numărul total de observații dintr-o anumită boală se presupune a fi mai mare trebuie să satisfacem legea numerelor mari, cercetînd un număr de 100—200 de cazuri din boala respectivă, pentru a putea trage concluzi generale.*

Ținînd seama de faptul că îmbolnăvirile variază în mare măsură în funcție de sex, grupe de vîrstă, boli asociate, complicații, starea generală a organismului etc. vor trebui cercetate un număr și mai mare de cazuri, și anume cel puțin 30 din fiecare categorie de sex, vîrstă etc.

Posibilitatea limitată de a cerceta un număr mare de bolnavi poate fi compensată prin extinderea observației clinice la un număr mai mare de cercetători, prin extinderea colaborării.

O greutate în generalizarea concluziilor sau în comparabilitatea observațiilor clinice este utilizarea unor noțiuni vagi, descriptive în ceea ce privește simptomatologia sau evoluția bolii, rezultatul tratamentului etc., în locul exprimării cantitative cifrice. Termeni generali ca: „dureri surde", bolnavul „se simte mai bine", bolnavul „acuză dureri mai mari" etc., nu numai că nu pot servi ca etalon pentru comparabilitatea și generalizarea observațiilor, dar nici pentru cunoașterea ulterioară a cazului observat.

Este, deci, necesar, chiar și pînă la descoperirea de noi metode de apreciere cantitativă a fenomenelor fiziologice ale organismului uman,

să se utilizeze, în măsură maximă posibilă, date cifrice privind aspectele calitative (de exemplu : în locul termenilor, ca : „dureri mari" se utilizează localizarea precisă a durerii, periodicitatea, iradierea durerii, durata ei pe grupe de bolnavi cu manifestări diferite etc.).

În continuare, redăm cele mai importante metode de statistică matematică ce pot fi utilizate în observaţia clinică, putînd fi completate cu elementele cuprinse în capitolul I.

2. CALCULUL VALORII MEDII, AL GRADULUI DE OMOGENITATE ŞI VARIABILITATE. EROAREA MEDIE A VALORII MEDII. CERTITUDINEA DIFERENŢEI DINTRE DOUĂ SAU MAI MULTE REZULTATE

În decursul cercetărilor clinice se concretizează numeroase observaţii privind debutul bolilor, evoluţia şi durata acestora, vindecarea sau decesul care survine la finele bolii, complicaţiile, evoluţia în cazul bolilor asociate etc. Se acumulează, astfel, o serie de date de către diferiţi cercetători, care cu toată variabilitatea lor au unele caracteristici esenţiale comune. Utilizarea metodelor matematice permite deducerea acestor valori medii şi gradul de dispersiune a cazurilor individuale de la această valoare medie. De asemenea, statistica permite cunoaşterea gradului de eroare a valorilor medii, obţinute pe numărul de cazuri cercetate faţă de valoarea reală (ideală, absolută) ce s-ar fi obţinut în cazul cercetării totalităţii cazurilor din grupa de boală sau boala respectivă.

În cercetările dinamice sau în cele comparative între diferite observaţii clinice distincte, statistica poate releva, în cazul diferenţelor dintre mai multe rezultate obţinute pentru aceeaşi categorie de bolnavi, dacă această diferenţă este reală sau doar întîmplătoare.

Să luăm un exemplu : în Spitalul de boli contagioase A., în cursul anului 1960, au decedat 30 din 282 de bolnavi de poliomielită. În anul 1961, în acelaşi spital, au decedat doar 18 cazuri dintr-un număr identic de bolnavi de poliomielită. Cum indicele de decese la suta de bolnavi $\left(\frac{\text{numărul deceselor. 100}}{\text{numărul bolnavilor}}\right)$ este în scădere în anul 1961 faţă de anul 1960, se pune întrebarea dacă această scădere indică o îmbunătăţire reală sau este vorba doar de o scădere întîmplătoare, datorită grupelor neomogene (ca vîrstă, sex, starea generală a organismului, complicaţii, forma de boală etc.) sau a numărului scăzut de cazuri de boală.

Presupunem că cele două grupe de bolnavi nu se deosebesc, în general, între ele, adică există două eşantioane omogene. Calculăm media, dispersiunea de la medie, eroarea medie şi eroarea diferenţei, care ne arătă certitudinea diferenţei.

Indicii medii în cei doi ani se calculează după formula anterioară : $\frac{\text{numărul deceselor. 100}}{\text{numărul bolnavilor}}$. Vom obţine astfel :

$$\frac{30 \cdot 100}{282} = 10{,}64\% \qquad \frac{18 \cdot 100}{282} = 6{,}38\%$$

Pentru calculul deviației-tip se utilizează formula utilizată pentru frecvențe :

$$\sigma_1 = \pm \sqrt{\frac{P_1 \cdot (100 - P_1)}{N_1}} = \pm \sqrt{\frac{10{,}64 \cdot 89{,}36}{282}} = \pm \sqrt{3{,}37} = \pm 1{,}84\,;$$

$$\sigma_2 = \pm \sqrt{\frac{P_2 \cdot (100 - P_2)}{N_2}} = \pm \sqrt{\frac{6{,}38 \cdot 93{,}62}{282}} = \pm \sqrt{2{,}12} = \pm 1{,}45.$$

Calculăm eroarea medie a celor doi indici de decese la %.

$$eM_1 = \pm \frac{\sigma_1}{\sqrt{N_1}} = \pm \frac{1{,}84}{\sqrt{282}} = \pm 0{,}11\,;$$

$$eM_2 = \pm \frac{\sigma_2}{\sqrt{N_2}} = \pm \frac{1{,}45}{\sqrt{282}} = \pm 0{,}09.$$

Ca să stabilim în ce măsură indicii de decese la 100 de bolnavi caracterizează grupele de bolnavi, calculăm procentul erorii medii :

$$eM_1\% = \pm \frac{eM_1 \cdot 100}{M_{p_1}} = \pm \frac{0{,}11 \cdot 100}{10{,}64} = \pm 1{,}03\,\%$$

$$eM_2\% = \pm \frac{eM_2 \cdot 100}{M_{p_2}} = \pm \frac{0{,}09 \cdot 100}{6{,}38} = \pm 1{,}41\,\%$$

Putem constata că proporția erorii medii nu depășește 3% la cei doi indici, deși fiind peste 1% este relativ mare.

Aplicînd triplul erorii medii putem să ne dăm seama de gradul de dispersiune a indicelui față de valoarea reală (absolută, ideală) :

$$3\,eM_1 = 3 \cdot (\pm 0{,}11) = \pm 0{,}33\,;$$

$$3\,eM_2 = 3 \cdot (\pm 0{,}09) = \pm 0{,}27.$$

Deci valoarea reală a celor doi indici se va găsi destul de apropiată de valoarea indicilor noștri, calculați pe un număr relativ scăzut de cazuri :

$$10{,}64\,\% \pm 0{,}33 = 10{,}31\,\% \text{ și } 10{,}97\,\%\,;$$

$$6{,}38\,\% \pm 0{,}27 = \ \ 6{,}11\,\% \text{ și } \ 6{,}65\,\%.$$

Din cele anterior expuse rezultă că fiecare dintre cele două grupe de bolnavi formează eșantioane omogene, avînd indici medii reprezentativi.

Pentru a ne da seama și de certitudinea diferenței celor doi indici, deci pentru a putea trage concluzii că în anul 1961 a scăzut în mod real indicele de decese la suta de bolnavi, calculăm eroarea diferenței :

$$eD = \pm \sqrt{eM_1^2 + eM_2^2} = \pm \sqrt{0{,}11^2 + 0{,}09^2} = \pm 0{,}14.$$

Triplul erorii diferenței $3 \cdot (\pm 0{,}14) = \pm 0{,}42$ este mult mai scăzut decît diferența dintre cei doi indici de decese la % (10,64% − 6,38% = = 4,26%), de unde se poate trage concluzia (cu o probabilitate de 99,73%) că în anul 1961 scăderea indicelui n-a fost întîmplătoare, ci certă.

Acest calcul privind certitudinea diferenței poate fi mult ușurat prin utilizarea unor tabele speciale, cu ajutorul cărora se poate afla rapid certitudinea diferenței dintre doi indici.

Tabelele apărute în Editura Academiei de științe medicale din U.R.S.S. sînt cele mai corespunzătoare din acest punct de vedere (tabelul XII).

În prima coloană căutăm cifra cea mai apropiată ce corespunde indicelui mai mic de decese la suta de bolnavi, în cazul nostru 6,38%. Vom lua, deci, cifra 6. Pe coloanele orizontale căutăm cifra cea mai apropiată de valoarea eșantionului nostru ce corespunde numărului de cazuri observate din fiecare an și anume 282. Vom lua cifra 300. La intersecția rîndului orizontal cu coloana verticală, citim din tabel valoarea 4,5, apropiată de valoarea diferenței dintre indicii de decese la suta de bolnavi, care a fost de 4,26%, ceea ce denotă că există o certitudine în ceea ce privește diferența dintre cei doi indici, ținînd seama și de faptul că la calculul anterior privind triplul erorii diferenței am obținut o valoare deosebit de mică (0,42), față de diferența mare dintre cei doi indici (4,26).

Se pot da numeroase exemple de acest fel. De exemplu: scăderea duratei medii de spitalizare în bronhopneumonie, pneumonie, în funcție de depistarea și internarea precoce a cazurilor, scăderea complicațiilor postoperatorii în rezecția gastrică sau gastroenteroanastomoză, în funcție de diferite tehnici operatorii utilizate, reducerea complicațiilor în diferite boli prin utilizarea diverselor medicamente etc. În toate aceste cazuri utilizarea indicilor anterior descriși confirmă sau infirmă certitudinea diferențelor, omogenitatea eșantioanelor luate în studiu.

3. CALCULUL DISPERSIUNII ÎN CAZUL EVENIMENTELOR RARE

În cazul observațiilor clinice nu totdeauna sîntem în măsură să alegem eșantioane reprezentative pentru a putea trage concluzii certe în privința omogenității sau a gradului de dispersiune.

De exemplu, nu vom putea lua un număr suficient de cazuri de febră butunoasă, *Coxsackie* etc., pentru a trage concluzii certe în privința simptomatologiei, deoarece numărul cazurilor cunoscute este deosebit de scăzut.

De asemenea, n-am putea găsi un eșantion reprezentativ pentru cunoașterea mortalității prin scarlatină, numărul cazurilor de mortalitate prin scarlatină fiind infim, în țara noastră.

Alteori mortalitatea poate fi zero, deși la un număr mai mare de cazuri sau într-un alt eșantion ar putea să survină cu certitudine cazuri mortale.

În astfel de situații, cînd nu putem recurge la eșantioane reprezentative, utilizăm alte metode de calcul, și anume, în locul distribuției normale distribuția Poisson, sau calculul după graficul din fig. 7.

Să luăm un exemplu în care evenimentul este egal cu zero.

Astfel, decesele la suta de bolnavi, pe un număr de 18 cazuri de encefalită epidemică, au fost zero, deși ne-am fi așteptat, în cazul unui număr mai mare de bolnavi, la o mortalitate destul de crescută.

Care ar putea să fie indicele de decese în cazul unui număr mai mare de cazuri, și anume, care ar fi limita superioară a acestei mortalități?

Pentru aceasta utilizăm scara pentru aplicarea limitei superioare a probabilității (99,73%, corespunzătoare lui 3σ) din graficul din fig. 7.

La o frecvență (n) de 18 cazuri îi corespunde pe grafic un $p_{0,n}$ egal cu 5,1.

Formula după care calculăm limita superioară probabilă a mortalității este :

$$m = \frac{p_{0,n}}{n} = \frac{5,1}{18} = 0,2833\,;$$

$$m\% = 100 \cdot 0,2833 = 28,33\%.$$

Acest rezultat arată că indicele deceselor la 100 de bolnavi în encefalita epidemică nu este zero decît întîmplător şi că, la un alt eşantion şi în orice caz la un eşantion reprezentativ, indicele de decese la suta de bolnavi ar fi destul de mare.

Se poate afirma cu o probabilitate de 99,73%, deci cu o mare certitudine, că limita maximă a acestei mortalități n-ar depăşi, în nici un caz, 28,33%, aşa cum ne arată rezultatul formulei.

Aceeaşi metodă se poate utiliza şi în cazul în care decesele printr-o anumită boală, de exemplu tetanos, ar fi fost 100% la un număr redus însă de cazuri (15 cazuri), care s-au soldat toate cu deces.

Citim pe aceeaşi scară valoarea $p_{0,n}$, ce corespunde unei frecvențe de 15. Aceasta este de 4,9.

Pe baza formulei obținem :

$$m = \frac{p_{0,n}}{n} = \frac{4,9}{15} = 0,3267\,;\ m\% = 100 \cdot 0,3267 = 32,67\%.$$

Se poate, deci, afirma, cu o probabilitate de 99,73% (deci cu o probabilitate de eroare doar de 0,27%), că la un număr mai mare de cazuri de tetanos, decesele la suta de bolnavi n-ar fi fost 100%, ci mult mai scăzut, limita maximă nedepăşind 32,67%.

4. CALCULUL DIFERENȚEI ÎN CAZUL FRECVENȚELOR MICI

În cazul frecvențelor mici sub 30, dar nu mai mici de 5, se utilizează pentru calcularea diferenței dintre două serii de variație, nu dispersiunea utilizată în distribuția normală, ci metoda χ^2 (chi pătrat), pe baza tabelei cvadruple.

Acest procedeu (χ^2) simplifică foarte mult calculele, cum s-a arătat şi în capitolul I, nemaifiind necesar de a se calcula deviația standard, coeficientul de variație, eroarea medie, coeficientul erorii medii, eroarea medie a diferenței etc.

Să luăm exemplul mortalității în bronhopneumonie la un număr de 20 de copii tratați cu penicilină, comparativ cu un lot identic de copii tratați cu aureomicină. În primul caz au avut loc 6 decese, în al doilea caz, 4 decese.

Se pune problema dacă această diferență de mortalitate este sau nu certă, dat fiind numărul scăzut de cazuri observate.

Pentru calcularea lui χ^2 alcătuim tabela cvadruplă (tabelul XVII).

Tabelul XVII

Procedeul χ^2 pentru calculul diferenței dintre două serii de variație

	Au decedat	N-au decedat	Total
Bolnavi tratați cu penicilină	$6_{(a_1)}$	$14_{(b_1)}$	20 $(a_1 + b_1 = n_1)$
Bolnavi tratați cu aureomicină	$4_{(a_2)}$	$16_{(b_2)}$	20 $(a_2 + b_2 = n_2)$
Total	10 $(a_1 + a_2 = n_3)$	30 $(b_1 + b_2 = n_4)$	40 $(n_1+n_2=n_3+n_4=N)$

Formula cunoscută a lui χ^2 este următoarea :

$$\chi^2 = \frac{(a_1 \cdot b_2 - b_1 \cdot a_2)^2}{n_1 \cdot n_2 \cdot n_3 \cdot n_4} \cdot (N - 1);$$

$$\chi^2 = \frac{(6 \cdot 16 - 14 \cdot 4)^2}{20 \cdot 20 \cdot 10 \cdot 30} \cdot (40-1) = 0{,}52.$$

Știm că în cazul în care χ^2 este mai mare decît 9, certitudinea diferenței dintre cele două serii este asigurată cu o probabilitate de 99,73%; în cazul în care χ^2 este mai mare decît 6,64 certitudinea diferenței între cele două serii este asigurată cu o probabilitate de 99,0%, iar atunci cînd depășește 3,84, doar cu o probabilitate de 95,0%.

În cazul nostru χ^2 este mult sub această valoare, astfel că diferența dintre cele două serii de variație, respectiv dintre mortalitatea cazurilor de bronhopneumonie, tratate cu penicilină și aureomicină, nu este certă.

Să luăm un alt exemplu similar, presupunînd 30 de cazuri de bronhopneumonie tratate cu penicilină cu 8 decese și 37 tratate cu aureomicină cu 2 decese.

Să construim tabela cvadruplă (tabelul XVIII).

Tabelul XVIII

Procedeul χ^2 pentru calculul diferenței dintre două serii de variație

	Au decedat	N-au decedat	Total
Bolnavi tratați cu penicilină	8	22	30
Bolnavi tratați cu aureomicină	2	35	37
Total	10	57	67

Aplicăm formula χ^2 :

$$\chi^2 = \frac{(8 \cdot 35 - 22 \cdot 2)^2}{30 \cdot 37 \cdot 10 \cdot 57} \cdot (67 - 1) = 5{,}8.$$

În acest caz putem aprecia cu o probabilitate de peste 95,0 % că diferența dintre decesele la cele două loturi de copii bolnavi de bronhopneumonie, tratați unii cu penicilină și ceilalți cu aureomicină, este certă.

Reamintim faptul că cele două eșantioane trebuie să fie omogene, în cazul nostru de același sex, vîrstă, aceeași zi a depistării bolii etc., căci altfel riscăm să greșim în concluziile noastre.

5. CALCULUL EȘANTIONULUI REPREZENTATIV ÎN OBSERVAȚIILE CLINICE

O problemă deosebit de importantă ce se pune și în domeniul cercetării clinice este problema eșantionului reprezentativ. Se pot trage concluzii generale pe baza unui eșantion, cînd se cunoaște că acesta reprezintă o parte din numărul existent sau posibil de cazuri.

Să reluăm exemplul privind scăderea indicelui de decese la suta de bolnavi prin poliomielită și să vedem dacă numărul de 282 de cazuri observate în fiecare an este suficient pentru a se putea trage concluzii generale sau nu.

Utilizăm formulele deja cunoscute :

$$\sigma D = \pm \sqrt{\frac{P_1 \cdot (100 - P_1) + P_2 \cdot (100 - P_2)}{N}};$$

$$N = \frac{P_1(100 - P_1) + P_2(100 - P_2)}{\sigma D^2}$$

Știind că $3\sigma D$ nu trebuie să depășească diferența dintre cei doi indici (10,64 %—6,38 % = 4,26 %), înseamnă că va trebui să luăm un eșantion a cărei eroare medie triplă a diferenței să nu depășească 4,26, deci să fie de 1,41, căci triplul ei va fi de 4,23, sub valoarea lui 4,26.

Completînd cu cifre simbolurile din formula anterioară vom găsi mărimea eșantionului reprezentativ :

$$N = \frac{10{,}64 \cdot 89{,}36 + 6{,}38 \cdot 93{,}62}{1{,}41^2} = 778.$$

Rezultă că ar trebui să avem, nu 282, ci 778 cazuri în observație pentru a putea trage concluzii general-valabile privind scăderea indicelui de decese la suta de bolnavi de poliomielită (cu o probabilitate de 99,73 %).

În acest caz, dacă am fi căutat în tabelul XII certitudinea diferenței celor doi indici de decese la % (10,64 și 6,38) la intersecția rîndului ce corespunde cifrei 6 cu coloana ce corspunde eșantionului 800 (eșantionul reprezentativ 778 rotunjit la sută este de 800), am fi găsit valoarea de 2,6 mult mai scăzută decît diferența dintre cei doi indici (4,26). A-

ceasta înseamnă că la un eșantion reprezentativ de 778 de cazuri de poliomielită scăderea indicelui de decese la suta de bolnavi în anul 1961 ar fi fost demonstrată cu o mare certitudine.

Dacă ne mulțumim și cu o probabilitate de 95,45% (2σ), ceea ce în cazul cercetărilor clinice este suficient, σD ar putea fi luat egal cu 2,13 (care înmulțit cu 2 dă 4,26, exact cît diferența dintre cei doi indici 10,64% — 6,38% = 4,26%).

În acest caz, în formula anterioară înlocuind σD^2 cu $2{,}13^2$ în loc de $1{,}41^2$ vom obține eșantionul reprezentativ egal cu 341 de cazuri de poliomielită, necesare a fi observate pentru a obține concluzii general-valabile (cu o probabilitate de 95,45%).

În această situație căutînd în tabelul XII, certitudinea eșantionului vom găsi la întretăierea rîndului 6 cu coloana 300 (cifra 341 fiind ceva mai apropiată de 300 decît de 400), indicele 4,5 cu puțin mai mare decît diferența dintre cei doi indici (4,26%). Deci și din tabel rezultă că eșantionul de 341 de cazuri ar fi suficient pentru concluziile generale.

Calculul anterior al mărimii eșantionului reprezentativ este destul de greoi. Tabelul XVI ne ușurează și mai mult în acest calcul.

Să luăm un exemplu: utilizînd un tratament pur simptomatic în cazul îmbolnăvirii prin febră tifoidă, mortalitatea observată în condiții obișnuite (pe grupe omogene) ar fi de 6,0%.

Utilizînd nitrofuranul pe un număr identic de bolnavi, mortalitatea ar fi doar de 2,0%.

Se cere să se afle în ce măsură acest indice de mortalitate de 2,0%, la bolnavii în cauză, este sau nu reprezentativ, suficient pentru a trage concluziile arătate. Diferența dintre cei doi indici este de 4,0% (6,0% — 2,0% = 4,0%).

Pentru a fi siguri că diferența dintre cei doi indici este certă trebuie să admitem, ca eroare medie dublă a celor 2 indici de letalitate, un maxim de $\pm$ 2,0%, căci altfel s-ar putea presupune o diferență doar întîmplătoare a acestor indici, deci superioritatea tratamentului prin nitrofuran n-ar fi demonstrată în mod cert.

Citim în tabelul XVI valoarea ce corespunde intersecției dintre rîndul orizontal 2 (letalitatea în cazul aplicării nitrofuranului) și coloana verticală 2 (maximul admis ca eroare a mediei).

Vom găsi în tabel valoarea 200, ceea ce înseamnă că este necesar un număr de 200 bolnavi tratați cu nitrofuran pentru a putea trage concluzia eficienței în terapeutica febrei tifoide. Deci putem afirma cu o probabilitate de peste 95,45% superioritatea tratamentului.

Pentru o certitudine și mai mare (99,73%) (deci eroarea medie triplă) trebuie să luăm aproximativ dublul acestei valori, deci 400 de bolnavi tratați pentru a putea trage concluzii absolut certe, privind eficiența tratamentului cu nitrofuran în febra tifoidă.

CAPITOLUL AL IV-LEA

METODE MATEMATICE ÎN CERCETĂRILE STĂRII DE SĂNĂTATE A POPULAȚIEI

1. METODE DE CERCETARE DEMOGRAFICĂ

PARTEA GENERALĂ

Cercetările privind numărul, structura sau dinamica populației se pot efectua datorită înregistrărilor curente ale datelor de stare civilă și a recensămintelor.

Ținînd seamă că aceste metode pot cuprinde o serie de observații asupra întregii populații, ele constituie o deosebită valoare.

Înregistrările curente și recensămintele, executate în mod științific, constituie o metodă de bază în studiul populației.

Sînt, însă unele situații, cînd aceste înregistrări totale nu se efectuează, cum este cazul într-o serie de țări din Africa sau Asia, unde organele oficiale ale stării civile se mărginesc la simple înregistrări în orașe sau localități rurale populate, dar omit marea masă a populației, chiar cu ocazia efectuării recensămintelor.

Astfel, după datele oficiale ale O.N.U.[1] rezultă că din 214 țări existente în lume doar, în 58 de țări se poate vorbi de o statistică demografică completă sau aproape completă, înțelegîndu-se prin aceasta o înregistrare de peste 90%. În alte 56 de țări, datele demografice sînt considerate incomplete, în 54 de țări gradul de completare nu este cunoscut, iar restul de 46 de țări, în majoritate din Africa (17 țări) și Asia (20 de țări) nu trimit de loc date demografice la O.N.U.

În astfel de situații, estimarea datelor privind numărul, structura și dinamica populației devine o sarcină dificilă, ducînd, de cele mai multe ori, la erori importante.

Totuși și în aceste cazuri, există metode, care, aplicate în mod corect, pot duce la rezultate cu o probabilitate de eroare ce nu alterează prea mult esența cercetării, dîndu-ne astfel relații importante.

---

[1] Annuaire démographique, 1956, New-York, pag. 77.

Astfel o metodă de selecție, just utilizată, pe un eșantion reprezentativ permite să se obțină unele concluzii privind populația, atît în ceea ce privește numărul și structura cît și dinamica ei.

În țările în care recensămintele se efectuează în mod periodic, din 10 în 10 ani și pe întreaga populație sau unde înregistrările de la nivelul birourilor de stare civilă (înregistrarea născuților vii și morți, a deceselor, căsătoriilor, divorțurilor, a migrației) se face cu multă precizie, se obțin, prin prelucrarea datelor de evidență statistică, o serie de date de precizie mult mai mare decît în cazul metodelor de selecție.

Cu toate acestea, înregistrările și cercetările asupra întregii populații nu pot aprofunda toate aspectele problemelor complexe ale acesteia. Unele probleme, cum ar fi numărul, structura pe sexe, grupe de vîrstă, starea civilă, profesiunea ca și cele privind dinamica populației pot fi prelucrate pe baza datelor recensămintelor sau ultimele pe baza înregistrărilor curente, dar o serie de aspecte necesare unui studiu aprofundat nu pot fi obținute pe această cale. Din această cauză este nevoie de metode adjuvante de înregistrare curentă sau periodică a unor date pe eșantioane reprezentative.

Astfel, datele obținute prin prelucrarea buletinelor de deces nu sînt suficiente în cazul mortalității infantile pe sexe, grupe de vîrstă și cauze, corelată cu starea social-economică, gradul de instruire a părinților, asistența acordată de către cadrele medico-sanitare, data decesului față de debutul bolii și față de data internării în spital etc. Aceste probleme sînt însă deosebit de importante pentru cunoașterea aprofundată a cauzelor complexe care au dus la deces, precum și în cunoașterea eficienței activității cadrelor medico-sanitare.

Sînt necesare deci, observații complementare, care nu trebuie neapărat să cuprindă întreaga colectivitate, ci doar un eșantion reprezentativ.

Anchetele reprezintă forme importante de cercetare parțială, care pot aprofunda o serie de probleme complexe din sînul populației, ce nu se pot obține pe baza datelor de recensămînt sau pe baza înregistrării curente a datelor de stare civilă. Astfel se completează observația totală cu cea parțială.

Legea numerelor mari ne arată că, dacă numărul total de observații este mai mare, mai apropiat de „univers", probabilitatea erorii este mai mică și invers, bineînțeles cu condiția de a cerceta colectivități omogene.

Pe de altă parte, extinderea cercetării la un număr mare de subiecte observate se face, de cele mai multe ori, în detrimentul omogenității.

S-a mai arătat anterior că legea numerelor mari nu este o lege fundamentală, un fetiș bun pentru rezolvarea tuturor problemelor, ci doar o metodă bună în cazul în care analiza calitativă a „esenței materiale" a fenomenelor a dovedit verosimilitatea problemelor cercetate. Iată de ce cercetătorul, care trebuie să facă o sinteză din observațiile diferitelor procese demografice, nu trebuie să rămînă străin de însăși cercetarea care se efectuează; din contră, să inițieze el însuși metodologia înregistrării și prelucrării materialului, pentru a se putea trage concluzii juste.

Înregistrarea totală, continuă, cum ar fi cea a datelor stării civile privind mișcarea naturală a populației sau periodică, cum ar fi recensă-

mîntul populației, sînt efectute de mii de persoane. Tocmai din aceste motive este necesar ca aceste înregistrări totale să se rezume doar la aspecte generale sau aspecte care nu pot da loc la interpretări eronate. Ele trebuie să fie completate cu cercetări parțiale, selective sau monografice, făcute de un personal mai puțin numeros dar de specialitate. În acest fel cercetările monografice sau anchetele, aprofundînd toate aspectele complexe ale problemei pot prezenta, uneori, o valoare științifică mult superioară cercetării totale.

Din formula erorii medii a mediei aritmetice $\left(e\,M = \pm \frac{\sigma}{\sqrt{N}} \right)$ se știe că eroarea mediei este invers proporțională cu rădăcina pătrată a numărului total de cazuri observate, deci pentru a scădea eroarea mediei la jumătate nu este suficient să creștem numărul de cazuri observate la dublu, ci mult mai mult, după cum rezultă din exemplul următor :

Să luăm exemplul erorii medii la înălțimea nou-născuților de sex masculin.

$$e\,M = \pm \frac{\sigma}{\sqrt{N}} = \pm \frac{1{,}6}{\sqrt{258}} = \pm \frac{1{,}6}{16{,}1} = \pm 0{,}1 \text{ cm.}$$

Dacă am vrea să scădem eroarea mediei la jumătate, deci la $\pm 0{,}05$ cm, pentru a avea o mai mare precizie în stabilirca mediei aritmetice ponderate, nu este suficient să creștem la dublu numărul de cazuri, deci la 516 în loc de 258. Rădăcina pătrată din 516 este de 22,7 deci nu dublul lui 16,1 (obținut prin extragerea rădăcinii pătrate din 258). Pentru a obține dublul lni 16,1, mai bine zis dublul $\sqrt{258}$, trebuie să creștem numărul de cazuri cercetate nu de două ori, ci de circa patru ori, de la 258 la 1 037, căci $\sqrt{1\,037} = 32{,}2$.

Tot astfel, pentru a scădea eroarea medie a mediei de la $\pm 0{,}1$ cm la un sfert, adică la 0,025 cm nu ar fi suficient să creștem numărul de cazuri observate de la 258 la 1 032, adică de patru ori, ci la 4 146 de cazuri, adică de circa 16 ori.

În aceste exemple presupunem că σ rămîne constant.

De aici rezultă că în practică nu trebuie să creștem neapărat numărul de observații excesiv de mult, căci aceasta îngreuiază cercetarea, fiind în detrimentul omogenității, ci mai bine să cercetăm grupe omogene de cazuri, căci eroarea este direct proporțională cu deviația tip (σ), adică eroarea scade direct proporțional cu gradul de omogenitate exprimat prin σ.

Criteriul omogenității nu trebuie însă să ducă pe cercetător să scadă prea mult numărul de cazuri observate, deoarece, în acest fel, eșantionul nu mai este reprezentativ.

Ținînd seama, atît de omogenitate, cît și de mărimea eșantionului, bazați fiind pe analiza calitativă a materialului, se poate ajunge la rezultatele cele mai juste.

Metoda de selecție privind cercetările asupra populației este deci o metodă deosebit de importantă de cunoaștere, atît în cazul imposibilității utilizării observației totale, continue sau periodice, cît și pentru aprofundarea problemelor cunoscute în general prin observația totală.

Problema se pune de a se utiliza, în mod just, metoda de selecţie în cercetările demografice şi în general în studiul stării de sănătate a unei colectivităţi, pentru ca rezultatele să se poată generaliza şi asupra întregii populaţii dintr-un teritoriu anumit: raion, regiune sau chiar din ţara întreagă.

Cercetarea trebuie să înceapă prin *fixarea scopului urmărit şi studiul materialului documentar* ce se referă la problema cercetată. Pe baza acestui studiu prealabil se pot stabili, deja o serie de ipoteze care urmează să fie verificate prin cercetare.

Chiar de această primă etapă depinde, în mare măsură, reuşita cercetării, căci încă în această etapă trebuie fixate metodele de cercetare, tehnica observaţiei, înregistrării şi prelucrării materialului.

Este greşit să se pornească de la bun început cu idei preconcepute sau premise pe care să căutăm să le justificăm cu orice preţ, neglijînd regulile elementare ale metodologiei statistice, să căutăm „să rotunjim" cifrele rezultate din cercetare, să le grupăm în mod nejust, pentru a obţine rezultatul presupus, dar care nu este cel real, sau nu este concluzia esenţială, ci secundară, incidentală. Şi într-un caz, şi în altul, eroarea este evidentă, iar cercetarea este necorespunzătoare, nu numai din punctul de vedere al generalizării asupra întregii populaţii, dar chiar şi pentru eşantionul cercetat.

Din acest punct de vedere, dacă este greşit de a porni de la orice premise şi idei preconcepute, este la fel de greşit de a face o cercetare fără a avea premise juste de la care să se poată porni la cercetare şi a putea, deci, alege metodologia cea mai potrivită. Însăşi documentaţia care trebuie făcută în prealabil, ne duce la astfel de premise juste pe care cercetarea le va dovedi ulterior.

Este bine chiar să insistăm asupra unor probleme care ni se par logice, dar care nu se conturează suficient în mersul cercetării. Perseverarea este „diabolică" doar în cazul în care neglijăm ceea ce este esenţial de dragul unor concluzii neesenţiale, pe care vrem cu tot dinadinsul să le demonstrăm. Perseverarea este însă necesară în cazul în care analiza calitativă a problemelor cercetate justifică aceasta.

Să luăm un exemplu din domeniul mortalităţii infantile. Din documentaţia existentă, bazată pe prelucrarea buletinelor de deces, cît şi pe baza cercetărilor parţiale, rezultă că din cei decedaţi sub vîrsta de 1 an aproximativ 28—30% au decedat sub vîrsta de o lună. Dacă vom alege un eşantion oarecare, s-ar putea întîmpla ca această proporţie de decedaţi sub o lună să fie mult mai mică, să zicem doar de 10—15%. Nu putem generaliza, în nici un caz, această observaţie, deoarece ne dăm seama că ea are doar un caracter accidental. Vom persevera chiar dacă după un nou eşantion cercetat vom ajunge la acelaşi rezultat şi, desigur, mărind eşantionul, extinzînd cercetarea şi în alte teritorii, vom putea constata că ne apropiem de proporţia aşteptată, adică de 28—30%

Desigur că în practică problema nu este atît de simplă. Sînt situaţii cînd premisele noastre nu au fost încă verificate de practică. În aceste cazuri situaţia este mai greu de clarificat şi va trebui ca prin continuarea cercetărilor să ajungem la rezultatele cele mai juste.

După fixarea scopului urmărit şi studiul documentaţiei existente, urmează delimitarea obiectului în timp, spaţiu şi volum.

Delimitarea obiectului cercetării în timp depinde de scopul urmărit şi de natura cercetării.

În studiul numărului, densităţii sau structurii populaţiei, intervalul de timp în care are loc cercetarea parţială nu trebuie să depăşească cîteva săptămîni, iar observaţia trebuie să se refere la un timp fix, numit moment critic, care trebuie să fie limitat la o anumită zi şi o anumită oră din ziua respectivă. Altfel survin schimbări care pot să altereze exactitatea rezultatelor.

Delimitarea obiectului în spaţiu şi volum constituie unul din aspectele importante ale metodei de selecţie. Se pune problema în cîte colectivităţi (comune, circumscripţii) trebuie efectuate observaţiile şi cît de mare trebuie să fie eşantionul ales? Şi aici natura cercetării şi scopul ei pot determina, atît numărul (cîte colectivităţi), cît şi mărimea eşantionului (numărul de cazuri).

În general 10% din întreaga colectivitate studiată este un eşantion reprezentativ, uneori în probleme de o mai mare omogenitate ajungînd şi un eşantion de 2%. În alte probleme, din contră, trebuie să creştem eşantionul la 20%.

Cu cît colectivitatea cercetată este mai omogenă, cu atît eşantionul poate fi mai mic, atît ca număr de colectivităţi, cît şi ca număr global de cazuri cercetate şi invers.

În cazul cercetării structurii populaţiei pe sexe sau pe grupe de vîrstă într-o regiune, este suficient să alegem doar cîteva eşantioane din diferite raioane şi circumscripţii, cu un număr total de cazuri care nu este nevoie să depăşească 5% din populaţia regiunii, în funcţie de caracteristica acestei probleme care are un grad destul de mare de omogenitate. Dacă populaţia regiunii este de 700 000 de locuitori, este suficient, deci, să cercetăm cîteva eşantioane cu un număr total de 35 000 de locuitori.

În cazul studiului dinamicii populaţiei: natalitatea, mortalitatea generală, mortalitatea infantilă, dată fiind omogenitatea mai mică decît în cazul structurii populaţiei pe sexe şi grupe de vîrstă, cercetarea trebuie să cuprindă 10—15%. Procentele arătate se referă la totalitatea problemei respective, numărul de decese generale sau decese sub 1 an.

Să luăm un exemplu concret. Populaţia regiunii A. este de 700 000 de locuitori. Să considerăm natalitatea de 20,0‰, mortalitatea generală de 10,0‰, iar cea infantilă de 6,0%. Cercetarea are ca scop să aprofundeze aceste fenomene demografice pe o serie de aspecte complexe ce nu rezultă din datele statistice oficiale. Care va fi numărul total al cazurilor şi din cîte colectivităţi trebuie alese aceste cazuri ce vor fi studiate, în vederea concluziilor pe care vrem să le generalizăm pe întreaga regiune?

Cunoaştem că fenomenele demografice sînt influenţate de starea social-economică a populaţiei în primul rînd, dar că ele pot fi influenţate, în mare măsură, şi de alţi factori cum ar fi: structura populaţiei pe sex şi grupe de vîrstă, ştiinţa de carte, condiţii igieno-sanitare, gradul de cultură sanitară, accesibilitatea şi gradul de adresabilitate a populaţiei la asistenţă medicală, anumite obiceiuri etc. Ţinînd seama de aceste condiţii, care toate influenţează dinamica populaţiei, trebuie să efectuăm cercetarea ţinînd seama de toate caracteristicile arătate. Pe baza studiului prealabil, ne putem da seama din cîte colectivităţi de locuitori trebuie să alegem eşantioanele.

Să presupunem că în total ne-am fixat la 10 grupe de populaţie. Numărul eşantionului nu se alege în mod mecanic, ci numai pe bază de documentare temeinică şi prealabilă, cînd ne convingem că aceste grupe de populaţie (din diferite teritorii sau de diferite ocupaţii, ştiinţă de carte etc.), se comportă diferit în ceea ce priveşte natalitatea, mortalitatea generală sau infantilă.

Mărimea eşantionului (numărul de cazuri observate), pentru fiecare grup în parte, nu trebuie să fie sub 30—40, dar este mai bine ca acesta să fie în jurul lui 100, ţinînd seama şi de omogenitatea mai mare sau mai mică a colectivităţilor.

În cazul nostru, ţinînd seama de cele 10 grupări făcute în cadrul populaţiei, vom avea, deci, de studiat în total 1 000 de născuţi vii, 1 000 de decese generale. În privinţa deceselor sub 1 an, ele constituind la noi mai puţin de o cincime din decesele generale, cele aproximativ 200 de cazuri n-ar fi suficiente căci grupate în 10 categorii, fiecare ar constitui un eşantion de 20 de cazuri, ceea ce n-ar putea fi reprezentativ. Trebuie, deci, să cercetăm cel puţin 400 de decese sub 1 an, pentru a avea, în acest caz, eşantioane de 40 de cazuri fiecare.

Să calculăm acum cît constituie aceste eşantioane din totalul cazurilor de născuţi vii şi decese generale şi sub 1 an din regiunea A.; dacă ele formează sau nu 10—15% în cazul născuţilor vii, al deceselor generale şi ale deceselor sub 1 an?

Am presupus o natalitate de 20‰, o mortalitate generală de 10‰ şi o mortalitate infantilă de 6,0% pe întreaga regiune A. În cifre absolute aceasta înseamnă, la o populaţie de 700 000 de locuitori, că anual se nasc 14 000 de copii, numărul de decese generale este de 7 000, iar cel al deceselor sub 1 an de 840.

Eşantionul în acest caz va cuprinde, deci, 7,14% din totalul născuţilor vii (1 000 din 14 000), în cazul mortalităţii generale 14,28% (1 000 din 7 100), iar în cazul mortalităţii infantile 47,62% (400 din 840).

Rezultă că în cazul natalităţii şi mortalităţii generale există un eşantion corespunzător, iar în cazul mortalităţii infantile un eşantion mare, mult peste necesar.

Desigur că aceste calcule sînt doar aproximative, ele verificîndu-se de-abia în urma prelucrării materialului cercetat prin calculele mediei aritmetice ponderate, a dispersiunii de la medie, a erorii medii a mediei etc.

Tehnica înregistrării parţiale în cercetările demografice

Se utilizează *metoda de înregistrare prin interogare directă*, aceasta ducînd la mai puţine erori de înregistrare, mai ales cînd cercetarea este făcută de specialişti cu un număr cît mai redus de echipe, avînd o coordonare unică.

Se mai poate utiliza *metoda autoînregistrării*, însă numai sub controlul personalului de specialitate, deşi chiar şi aşa pot surveni o serie de erori voluntare sau mai ales omisiuni, ceea ce alterează valoarea cercetării.

Cercetarea se poate face, fie prin *observare continuă sau curentă*, fie prin *observarea şi înregistrarea periodică* a problemelor. Metoda mai adecvată, utilizată rar de altfel, este cea mixtă: continuă şi periodică.

În cazul cercetării privind structura populaţiei, pe sexe, grupe de vîrstă, profesiuni, ştiinţă de carte (cînd nu s-au efectuat recensăminte sau cercetarea caută să adîncească datele obţinute prin recensămînt),

se utilizează, fie observaţia şi înregistrarea periodică, fie *înregistrarea de o singură dată*, cea periodică fiind, desigur, mult mai corespunzătoare unui studiu ştiinţific.

Este bine ca aceste cercetări parţiale, periodice, să se suprapună recensămintelor periodice, pentru a putea avea o comparabilitate pe întreaga colectivitate.

Cînd se efectuează cercetarea parţială a dinamicii populaţiei se utilizează observaţia şi înregistrarea curentă a fenomenelor pe un interval de aproximativ 5 ani, pentru a putea urmări problema evolutiv an de an, pentru a putea elimina interpretarea eronată a fluctuaţiilor posibile de la un an la altul a acestor fenomene.

Formele cercetării. Formele mai importante ale cercetării prin observaţie parţială sînt : monografia şi ancheta.

Monografia. Deşi se limitează ca spaţiu la un teritoriu mic : comună, circumscripţie sau cel mult un raion, monografia prezintă importanţă prin faptul că aprofundează problemele, datele fiind redate în mod amplu. Ţinînd seamă de complexitatea fenomenelor demografice, monografia poate fi forma cea mai adecvată de studiu pe un teritoriu limitat.

Ancheta statistică cuprinde un număr mult mai mic de probleme decît în cazul monografiei, de cele mai multe ori o singură problemă, în schimb se extinde la un spaţiu şi pe un interval de timp mult mai mare. Ancheta adînceşte problemele respective ; din aceste motive, chiar mai mult decît în cazul monografiei, ea trebuie efectuată de specialişti, în vederea unei mai mari omogenităţi ca metodă de muncă şi mai ales pentru analiza finală a materialului obţinut prin observaţie şi înregistrare.

Metoda de selecţie îşi are importanţa prin faptul că cercetările pot fi generalizate asupra întregii colectivităţi. Un eşantion de 2—15% din întreaga colectivitate a unui raion sau a unei regiuni (în funcţie de problema cercetată), dacă numărul colectivităţilor este bine ales poate fi reprezentativ pentru întreaga colectivitate.

Selecţia poate fi întîmplătoare, mecanică sau tipică.

Selecţia întîmplătoare se face, după cum arată şi numele, cu totul la întîmplare, fără a alege grupele pe criterii topografice, social-economice sau alte criterii.

Selecţia mecanică se face alegînd eşantioanele într-o anumită ordine, fiecare a 3-a — a 4-a circumscripţie sau fiecare a 10-a comună aleasă în ordine alfabetică sau ca tablele de şah. În realitate şi această metodă de selecţie este ca şi cea întîmplătoare, dar are o repartizare teritorială mai bună.

Selecţia tipică se face grupînd colectivităţile în categorii omogene, pe baza analizei iniţiale, după care se trece la selecţia mecanică.

În cazul cercetărilor demografice trebuie utilizată numai selecţia pe baza analizei calitative iniţiale, pe grupe cît mai omogen posibile, bineînţeles fără a alege numai ceea ce corespunde premiselor anterior fixate şi a elimina tot ceea ce nu corespunde acestor premise. Din aceste motive, gruparea trebuie să fie făcută în aşa fel încît să cuprindă eşantioane din toate categoriile distincte unele de altele şi în primul rînd, din colectivităţile care reprezintă o pondere mai mare. De exemplu, în cazul studiului natalităţii, în vederea cunoaşterii cauzelor denatalităţii, vom cerceta colectivităţi caracterizate printr-o natalitate scăzută sau foarte

scăzută. Desigur că va trebui să luăm eșantioane, grupe de colectivități la care natalitatea este crescută, altele cu natalitatea medie, nu numai ca eșantioane-martor, ca bază de comparare, dar și pentru studiul diferențial al acestei probleme. Numărul cel mai mare de eșantioane îl vom alege însă din colectivitățile cu natalitate scăzută.

În cazul mortalității generale și infantile, numărul cel mai mare de eșantioane îl vom lua din colectivitățile cu mortalitate mare, scopul fiind tocmai de a cerceta cauzele mortalității generale și infantile crescute. Pentru un studiu diferențiat și comparativ vom alege un număr mai mic de eșantioane și din colectivitățile cu mortalitate generală sau infantilă scăzută și medie, pentru a putea cunoaște mai bine cauzele complexe care contribuie la scăderea, menținerea sau creșterea mortalității.

Gruparea omogenă nu este numai o condiție în stabilirea inițială a eșantioanelor, ci și o condiție de grupare ulterioară pe caracteristici esențiale a materialului de cercetare, deja centralizat.

Gruparea pe caracteristici esențiale se face pe baza analizei calitative inițiale a studiului efectuat. Astfel în studiul mortalității infantile, gruparea pe sexe, grupe de vîrstă, cauze, ocupația părinților, gradul de cultură generală și sanitară, accesibilitatea și adresabilitatea la asistența medicală etc. constituie caracteristici esențiale. Materialul cercetat pe aceste caracteristici esențiale trebuie bineînțeles grupat în alte grupe caracteristice. Astfel, grupele de vîrstă nu se împart în mod mecanic pe zile, săptămîni sau luni, ci după felul cum anumite condiții complexe (social-economice, morbiditate etc.) determină o mortalitate mai mare în primele zile ale vieții copilului, în primele săptămîni sau în primele luni.

Cifre absolute, mărimi relative, mărimi medii. Rezultatele cercetării se cristalizează în cifre absolute, care după grupare vor fi aranjate într-o serie de tabele. Ca exactitate nimic nu poate fi mai util decît cifrele absolute. Pentru cei de specialitate un tabel alcătuit din cifre absolute constituie baza analizei finale a cercetării. Cu toate acestea, de cele mai multe ori, comparabilitatea dintre cifrele absolute devine greoaie, mai ales în cazul seriilor mari de variație sau în cazul cifrelor mari, de ordinul zecilor sau sutelor de mii, sau a zecimalelor. Din aceste motive se utilizează mărimi relative (procente, proporții), cele mai importante fiind: indicatorii extensivi, indicatorii intensivi și indicatorii intuitivi.

Indicatorii extensivi (de structură sau de repartiție). Un exemplu îl avem în cazul mortalității generale pe cauze, totalitatea cauzelor fiind considerată egală cu 100; grupele de cauze sau cauzele individuale constituie părți din acest total (procente) calculate după regula de trei simplă, suma lor totală fiind egală cu 100.

Astfel de indicatori extensivi sau structurali pot fi simpli, cum ar fi în cazul anterior, pot fi însă și complecși ca în cazul deceselor pe sexe, grupe de vîrstă și medii. În acest ultim caz fiecare dintre caracteristici va alcătui un procent din cadrul grupei respective, deci mediul rural plus urban va da 100%, sexul masculin și feminin de asemenea 100% etc.

Indicatorii intensivi (de frecvență) sînt proporțiile sau raporturile dintre o colectivitate față de o altă colectivitate luată ca bază de comparație; de exemplu proporția deceselor la o populație de 1 000 de locuitori

sau proporția deceselor sub 1 an, raportată la suta de născuți vii în decurs de 1 an.

Indicatorii intuitivi (demonstrativi) se utilizează în cadrul seriilor dinamice, una dintre mărimi fiind luată ca bază de comparabilitate și considerată egală cu 100, restul mărimilor calculîndu-se prin regula de trei simplă. Ca exemplu avem compararea natalității pe o perioadă de 10 ani, natalitatea în primul sau oricare dintre ani fiind luată ca indice de bază, considerat egal cu 100, natalitatea din ceilalți ani depășind sau fiind sub 100, după cum ea a fost mai mare sau mai mică decît natalitatea luată ca bază (100).

Există și alți indicatori, însă în cercetările demografice se utilizează de obicei cei anterior menționați.

Atît cifrele absolute, cît și cele relative nu pot servi ca bază de analiză, dacă ele nu sînt comparabile. Din această cauză se utilizează *mărimile medii*, care, calculate pe grupări omogene, devin într-adevăr, mărimi tipice pentru colectivitatea respectivă. Mărimile medii, pe grupări mari, neomogene, devin cifre medii globale, care ne induc în eroare, mascînd diferitele caracteristici ale fenomenelor cercetate, ștergînd toate asperitățile. În acest caz mărimea medie nu poate servi ca bază științifică de comparație pentru elementele grupei respective.

Este sarcina cercetătorului, ca pe baza documentației inițiale a problemelor în studiu, să grupeze materialul pe caracteristici esențiale pentru ca mărimea medie să poată prezenta o medie științifică. Dacă de exemplu, cercetarea privind mortalitatea infantilă nu va diferenția problema pe cauze, sex, grupă de vîrstă, condiții social-economice, cultural sanitare etc., ci se va limita doar la gruparea deceselor sub un an pe sexe sau pe grupe de vîrstă, concluziile nu vor fi edificatoare, în privința ansamblului de măsuri ce trebuie luate în vederea scăderii mortalității infantile.

Mărimile medii mai frecvent utilizate în cercetările demografice sînt media aritmetică simplă și ponderată, uneori și media progresivă.

Mărimea medie ne dă o idee generală despre o problemă, nu și despre o serie de valori, care se abat mai mult sau mai puțin de la această medie.

Două medii egale pot avea structuri diferite, una cu o dispersiune mare a elementelor (variantelor) de la medie, a doua cu o dispersiune mult mai mică de la medie. În aceste cazuri cele două colectivități nu sînt omogene, deși cu aceeași valoare medie, nu sînt totuși comparabile.

Iată cum mărimile medii ne pot duce, de multe ori la rezultate neștiințifice. Din aceste motive, în toate cercetările este necesar să se treacă la calculele dispersiunii de la medie : deviația — tip, coeficientul de variație, cît și la eroarea medie sau eroarea probabilă a mediei aritmetice ponderate.

Legăturile dintre caracteristicile unui fenomen pot fi redate concret prin calculul coeficientului de corelație și de regresie.

Toate aceste măsurători medii, măsurători ale dispersiunii de la medie, ale corelației și regresiei, cît și alte calcule expuse pe larg în capitolul I vor fi reluate, în continuare, și în acest capitol.

2. METODE DE CERCETARE DEMOGRAFICĂ

PARTEA SPECIALĂ

A. METODE DE CERCETARE PRIVIND NUMĂRUL ŞI STRUCTURA POPULAŢIEI

Nunărul şi structura populaţiei se poate cunoaşte cu ocazia recensămintelor. Problema se pune de a găsi metode de estimare a numărului şi structurii populaţiei, în cazul cînd nu se fac recensăminte sau în perioada dintre recensăminte.

Estimarea numărului de locuitori între două recensăminte, cînd nu există o înregistrare perfectă a dinamicii populaţiei şi mai ales a migraţiei, se poate face prin diferite metode.

Metoda interpolării se efectuează prin calcularea raţiei anuale de creştere. De exemplu, dacă numărul de locuitori într-o regiune a crescut cu 100 000, faţă de recensămîntul efectuat în urmă cu 10 ani, raţia anuală de creştere va fi de 10 000 (100 000 : 10). Avînd această raţie anuală putem să estimăm populaţia în orice an între cele două recensăminte. Mai mult, putem să estimăm cu această raţie anuală şi populaţia după ultimul recensămînt. În acest caz vorbim de *extrapolare*.

Calculul devine cu atît mai impreciş însă cu cît perioada dintre recensăminte este mai mare sau cu cît extrapolarea se face pe o perioadă mai îndepărtată de ultimul recensămînt (peste 3—4 ani).

O estimare mai precisă a populaţiei, în cazul în care migraţiile nu sînt înregistrate, se poate efectua prin metoda corectării raţiei anuale de creştere.

Corectarea raţiei anuale de creştere se face prin calcularea sporului natural al populaţiei, an de an între cele două recensăminte, stabilind procentul anual al sporului natural din totalul sporului dintre cele două recensăminte. Să presupunem că sporul natural al populaţiei din regiunea A., între două recensăminte ar fi de 120 000 faţă de 100 000, cît reprezintă sporul real care ţine seamă şi de migraţiuni. Calculăm an de an cît la sută reprezintă (din 120 000) sporul natural şi găsim, spre exemplu, următorii indici : 13%, 12%, 12%, 11%, 10%, 9%, 9%, 9%, 8%, şi 7%. Vom considera prin analogie şi raţia anuală de creştere în aceeaşi proporţie, calculînd, de data aceasta, nu la 120 000, ci la 100 000 locuitori, cît a fost diferenţa între cele două recensăminte. În acest caz, transformînd procentul anterior al raţiei anuale de creştere în cifre absolute, obţinem raţia anuală corectată : 13 000 ; 12 000 ; 12 000, 11 000 ; 10 000 ; 9 000 ; 9 000 ; 9 000 ; 8 000 şi 7 000. Utilizînd aceste valori în locul raţiei anuale necorectate (10 000), vom obţine o estimare mult mai precisă, pentru necesităţile extrapolării.

Estimarea structurii populaţiei pe sexe şi grupe de vîrstă între două recensăminte este mult mai dificilă. Ea se poate face cu exactitate în cazul cunoaşterii dinamicii populaţiei, inclusiv a migraţiei, calculînd an de an naşterile şi decesele pe sexe şi grupe de vîrstă ; la fel şi mişcarea mecanică a populaţiei.

Mai puţin exactă este calcularea structurii populaţiei pe sexe, grupe de vîrstă, ocupaţie etc., pe baza comparării celor două recensăminte, prin metoda interpolării şi extrapolării, cu raţia anuală de creştere. De

exemplu, dacă la recensămîntul anterior, grupa de vîrstă 0—14 ani a constituit 33% din totalul populației, iar la ultimul recensămînt, după 10 ani, doar 28%, putem estima că rația anuală de creştere a scăzut an de an cu 0,5% la această grupă de vîrstă.

Situația este mult mai grea, atunci cînd nu s-au efectuat de loc recensăminte generale ale populației. În acest caz cercetările parțiale, pe bază de eşantioane reprezentative, ne pot indica cu o oarecare aproximație, atît numărul, cît şi structura populației. Metoda de selecție se utilizează nu numai în aceste cazuri, ci chiar şi în cazul recensămintelor generale, pentru a verifica gradul de exactitate cu care ele au fost efectuate.

Să luăm următorul exemplu:

Pe baza unui eşantion cules din 18 raioane să estimăm proporția populației sub 5 ani din întreaga țară, considerînd că acest eşantion ar fi reprezentativ.

Sîntem ajutați în interpretarea justă a valorii eşantionului prin faptul că în realitate datele pe grupe de vîrstă le cunoaştem, atît pe țară cît şi pe regiuni şi raioane. Astfel vom putea constata cu precizie şi valabilitatea calculelor în alegerea eşantionului reprezentativ în stabilirea grupelor de vîrstă a populației.

Ne putem afla, însă, în situația de a nu cunoaşte de loc numărul şi structura populației dintr-un teritoriu, cercetarea avînd ca scop tocmai de a estima această structură, pe baza cercetării parțiale prin metoda de selecție. Desigur că în acest caz problema este cu mult mai dificilă, în sensul că lipseşte posibilitatea de a verifica exactitatea eşantionului. În cazul nostru, în scop instructiv, cunoaşterea numărului şi a structurii populației din întregul teritoriu al țării, cît şi din teritoriul celor 18 raioane pe care le considerăm drept eşantion, ne va uşura înțelegerea metodologiei cercetării selective şi a valorii ei, în generalizarea pe întreg teritoriul.

Dacă vom lua în cercetare, de exemplu, cîte 3 eşantioane din fiecare raion, vom obține un număr de 54 de eşantioane, cu un număr total de 1 800 000 locuitori. Populația sub 5 ani a eşantionului să presupunem că este de 180 000, ceea ce constituie 10,0% din eşantionul luat.

Ceea ce interesează în calculele noastre este dacă eşantionul total de 180 000 locuitori sub 5 ani este reprezentativ în ceea ce priveşte structura pe grupe de vîrstă sub 5 ani pentru întreaga populație a țării.

În acest scop vom calcula omogenitatea eşantionului (σ) şi mărimea reprezentativă a eşantionului (N), după formulele utilizate în capitolul I.

Fiind cazul de proporții, utilizăm următoarea formulă pentru deviația tip (σ):

$$\sigma = \pm \sqrt{\frac{P \cdot Q}{N}} = \pm \sqrt{\frac{P \cdot (100 - P)}{N}}.$$

Înlocuind cu cifre termenii din formulă vom obține:

$$\sigma = \pm \sqrt{\frac{10{,}0 \cdot 90{,}0}{1\,800\,000}} = \pm \sqrt{0{,}000\,5} = \pm 0{,}022.$$

Se constată că indicele deviației-tip este mic, ceea ce denotă o omogenitate destul de mare, fapt ce se exprimă mai sugestiv prin coeficientul de variație deosebit de mic.

$$cV = \pm \frac{\sigma \cdot 100}{Mp} = \pm \frac{0{,}022 \cdot 100}{10{,}0} = \pm 0{,}22\%.$$

Pentru a putea constata în ce măsură media de 10,0% exprimă într-adevăr proporția justă a copiilor sub 5 ani din totalul populației pe țară, cu alte cuvinte, cît de mare este eroarea mediei eșantionului cercetat față de media absolută, calculăm eroarea medie:

$$e\,M = \pm \frac{\sigma}{\sqrt{N}} = \pm \frac{0{,}022}{\sqrt{1\,800\,000}} = \pm \frac{0{,}022}{1\,341{,}6} = \pm 0{,}0000\,16.$$

Luînd chiar triplul erorii medii, indicele de eroare va fi, cu toate acestea foarte mic.

$$3 \cdot 0{,}0000\,16 = \pm 0{,}0000\,48.$$

Dacă transformăm în procente această eroare medie vom obține 0,000 16, ceea ce denotă că media aritmetică a eșantionului (10,0) exprimă în mod just, proporția pe țară a copiilor de 0—5 ani.

$$e\,M\% = \pm \frac{e\,M \cdot 100}{Mp} = \pm \frac{0{,}000016 \cdot 100}{10{,}0} = \pm 0{,}00016\%.$$

Luînd triplul procentului erorii medii obținem, de asemenea, o eroare mică:

$$3 \cdot 0{,}000\,16\% = \pm 0{,}00048\%.$$

În realitate proporția populației sub 5 ani pe întreaga țară, la recensămîntul din anul 1956 a fost de 10,57%, ceea ce înseamnă că media aritmetică obținută din eșantion prezintă o diferență mică față de situația reală pe țară, ceea ce confirmă reprezentativitatea eșantionului ales.

Metoda de estimare a exactității cercetării privind structura populației pe grupe de vîrstă. În cazul grupării populației pe vîrste, pentru verificarea exactității datelor înregistrate, mai bine zis a vîrstelor declarate de locuitori (în general înregistrările trebuie să fie efectuate pe bază de documente oficiale), se poate utiliza indicele de concentrație.

Indicele de concentrație se calculează în felul următor: în primul rînd, se adună numărul de locuitori avînd vîrsta cuprinsă între 23 și 62 de ani și se împarte la 5. În al doilea rînd: dintre acești locuitori (între 23 și 62 de ani) se însumează numărul de locuitori a căror vîrstă declarată se termină, fie cu 0, fie cu 5. În al treilea rînd: se împarte valoarea obținută din a doua însumare la valoarea obținută la prima, înmulțindu-se apoi cu 100 pentru a se obține procentul.

Indicele rezultat va avea, în general, o valoare cuprinsă între 100 și 500.

În cazul unui indice de concentrație cuprins între 100 și 105, putem considera că datele declarate și înregistrate privind vîrsta sînt foarte exacte. Un indice cuprins între 105 și 109,9 arată că înregistrarea vîrstelor a fost relativ exactă. Peste un indice de 110 se apreciază datele obținute

Se constată că indicele deviației-tip este mic, ceea ce denotă o omogenitate destul de mare, fapt ce se exprimă mai sugestiv prin coeficientul de variație deosebit de mic.

$$cV = \pm \frac{\sigma \cdot 100}{Mp} = \pm \frac{0{,}022 \cdot 100}{10{,}0} = \pm 0{,}22\%.$$

Pentru a putea constata în ce măsură media de 10,0% exprimă într-adevăr proporția justă a copiilor sub 5 ani din totalul populației pe țară, cu alte cuvinte, cît de mare este eroarea mediei eșantionului cercetat față de media absolută, calculăm eroarea medie:

$$e\,M = \pm \frac{\sigma}{\sqrt{N}} = \pm \frac{0{,}022}{\sqrt{1\,800\,000}} = \pm \frac{0{,}022}{1\,341{,}6} = \pm 0{,}0000\,16.$$

Luînd chiar triplul erorii medii, indicele de eroare va fi, cu toate acestea foarte mic.

$$3 \cdot 0{,}0000\,16 = \pm 0{,}0000\,48.$$

Dacă transformăm în procente această eroare medie vom obține 0,000 16, ceea ce denotă că media aritmetică a eșantionului (10,0) exprimă în mod just, proporția pe țară a copiilor de 0—5 ani.

$$e\,M\% = \pm \frac{e\,M \cdot 100}{Mp} = \pm \frac{0{,}000016 \cdot 100}{10{,}0} = \pm 0{,}00016\%.$$

Luînd triplul procentului erorii medii obținem, de asemenea, o eroare mică:

$$3 \cdot 0{,}000\,16\% = \pm 0{,}00048\%.$$

În realitate proporția populației sub 5 ani pe întreaga țară, la recensămîntul din anul 1956 a fost de 10,57%, ceea ce înseamnă că media aritmetică obținută din eșantion prezintă o diferență mică față de situația reală pe țară, ceea ce confirmă reprezentativitatea eșantionului ales.

Metoda de estimare a exactității cercetării privind structura populației pe grupe de vîrstă. În cazul grupării populației pe vîrste, pentru verificarea exactității datelor înregistrate, mai bine zis a vîrstelor declarate de locuitori (în general înregistrările trebuie să fie efectuate pe bază de documente oficiale), se poate utiliza indicele de concentrație.

Indicele de concentrație se calculează în felul următor: în primul rînd, se adună numărul de locuitori avînd vîrsta cuprinsă între 23 și 62 de ani și se împarte la 5. În al doilea rînd: dintre acești locuitori (între 23 și 62 de ani) se însumează numărul de locuitori a căror vîrstă declarată se termină, fie cu 0, fie cu 5. În al treilea rînd: se împarte valoarea obținută din a doua însumare la valoarea obținută la prima, înmulțindu-se apoi cu 100 pentru a se obține procentul.

Indicele rezultat va avea, în general, o valoare cuprinsă între 100 și 500.

În cazul unui indice de concentrație cuprins între 100 și 105, putem considera că datele declarate și înregistrate privind vîrsta sînt foarte exacte. Un indice cuprins între 105 și 109,9 arată că înregistrarea vîrstelor a fost relativ exactă. Peste un indice de 110 se apreciază datele obținute

ca fiind aproximative. În sfîrşit, un indice de concentraţie de peste 175 denotă inexactitatea datelor, ele nemaiputînd fi utilizate.

În general, gruparea vîrstei pe grupe din 5 în 5 ani duce la anihilarea nepotrivirilor, chiar şi în cazul unui indice ceva mai crescut de concentraţie. Totuşi, în cazul unui indice de concentraţie de peste 175, datele pe grupe de vîrstă nu mai prezintă vreo garanţie, chiar grupate pe intervale din 5 în 5 ani.

Desigur că în cazul recensămintelor populaţiei sau a cercetării de selecţie, lucrîndu-se cu echipe bine instruite, astfel de situaţii, adică înregistrări greşite, nu pot surveni.

De asemenea, specificăm că războaiele pot duce la schimbări importante în ceea ce priveşte structura unei populaţii pe grupe de vîrstă, ceea ce face ca indicele de concentraţie să nu poată fi utilizat. În măsură mai mică sau mai mare, indicele poate fi perturbat în cazul emigrărilor sau imigrărilor de populaţie, în special la grupele de vîrstă tinere, la care procesul de migraţie poate fi mai accentuat.

Să luăm un exemplu arbitrar pentru calculul indicelui de concentraţie : populaţia din regiunea A., pe ani de vîrstă (tabelul XIX).

Tabelul XIX

Populaţia regiunii A. (23—62 de ani)

Vîrsta	Numărul locuitorilor	Vîrsta	Numărul locuitorilor
23	21 520	43	17 405
24	22 408	44	16 055
25	24 638	45	18 786
26	22 458	46	15 915
27	21 019	47	14 625
28	21 460	48	15 387
29	19 238	49	14 114
30	22 575	50	14 955
31	18 433	51	12 327
32	20 360	52	13 233
33	20 367	53	12 480
34	17 512	54	10 898
35	19 385	55	14 522
36	16 315	56	12 059
37	9 558	57	8 561
38	8 231	58	9 296
39	8 407	59	9 009
40	11 615	60	11 239
41	13 685	61	7 908
42	17 173	62	8 962

Adunînd locuitorii de vîrstele cuprinse între 23 şi 62 de ani (X) şi împărţind suma la 5, adunînd, de asemenea, locuitorii de vîrstă care se termină cu zero şi cu 5 (Y), vom putea calcula indicele de concentraţie (C) după formula următoare :

$$C = \frac{\frac{Y}{X}}{5} \cdot 100 = \frac{137.715}{614.093} \cdot 100 = 112{,}1\%.$$

Acest indice de 112,1 % poate fi considerat ca fiind bun, ţinînd seama de consecinţele demografice ale celor două războaie mondiale care au dus la schimbări importante în structura populaţiei.

B. METODE DE CERCETARE PRIVIND DINAMICA POPULAŢIEI

În ţările unde se efectuează înregistrarea curentă a dinamicii populaţiei, pe bază de buletine care ulterior sînt prelucrate statistic, cunoaşterea fenomenelor demografice este posibilă în mod destul de amănunţit.

Problema se pune doar atunci cînd nu sîntem siguri de exactitatea înregistrărilor sau cînd înregistrările sînt doar parţiale, cum este cazul în cele mai multe ţări din Africa, Asia sau în unele ţări din America de Sud.

Dar nici înregistrările complete, datele demografice oficiale nu pot, de cele mai multe ori, satisface întru totul necesităţile de studiu complex şi din aceste motive trebuie utilizate şi în aceste cazuri cercetările parţiale.

Cele mai preţioase cercetări parţiale sînt cele de selecţie, monografiile sau anchetele demografice. Acestea nu sînt însă satisfăcătoare din toate punctele de vedere şi în special din punctul de vedere al generalizării rezultatelor. De aceea valoarea cercetărilor de selecţie este deosebit de importantă, bineînţeles doar în cazul în care eşantioanele sînt reprezentative pentru întreaga colectivitate din care au fost alese.

Buletinele oficiale de înregistrare la starea civilă nu sînt destul de cuprinzătoare şi din aceste motive se utilizează fişe sau buletine mult mai detaliate, după scopul urmărit.

Care sînt problemele ce se pun cercetătorului pentru a lămuri aspectele şi cauzele atît de complexe ale fenomenelor demografice?

În primul rînd, problemele mai importante sînt: cauzele scăderii natalităţii şi fecundităţii generale şi conjugale; cauzele mortalităţii generale şi infantile, pe medii, sexe, grupe de vîrstă; cauzele mortalităţii materne; ale mortinatalităţii; cauzele scăderii nupţialităţii sau ale creşterii divorţurilor; cauzele şi proporţia migraţiilor etc.

Desigur că cele mai multe dintre aceste probleme sînt atît de strîns legate una de cealaltă, încît studiul separat al mortalităţii generale, de exemplu, nu este posibil. Cauzele mortalităţii sînt strîns legate de structura populaţiei pe grupe de vîrstă, care, la rîndul ei, este influenţată de nivelul natalităţii, ca şi de mortalitatea la diferitele grupe de vîrstă. Natalitatea, la rîndul ei, este şi ea influenţată de nupţialitate, de divorţuri, de mortinatalitate, ca şi de structura populaţiei pe sexe şi grupe de vîrstă etc. Toate aceste fenomene sînt legate de proporţia şi sensul migraţiilor, de mortalitatea pe cauze şi grupe de vîrstă etc.

Desigur că dinamica populaţiei, ca şi structura ei sînt determinate, în primul rînd, de condiţiile social-economice, de structura social-economică a societăţii respective.

Legăturile acestea multiple sînt foarte complexe, influenţa reciprocă deosebit de mare, legăturile de la cauză la efect inversîndu-se şi întregindu-se.

Din aceste motive cercetările trebuie să cuprindă, uneori, întreg ansamblul de probleme, chiar şi în cercetările parţiale.

În cercetarea cauzelor care duc la scăderea natalității și a fecundității generale și conjugale, problemele mai importante care trebuie urmărite sînt : ocupația, venitul și în general bugetul familiei, numărul membrilor de familie, vîrsta părinților și a copiilor, condițiile de locuit, tensiunile în cadrul familiei, perspectivele pe care le are și mai ales cele pe care și le prevede familia respectivă în ceea ce privește condițiile de viață pe viitor. De asemenea, trebuie urmărite problemele de sănătate, în special anumite boli ca : sifilisul, gușa, tuberculoza, ginecopatiile care pot duce la sterilitate sau avorturi spontane, mortinatalitate etc.

În studiul cauzelor mortalității generale sau infantile, pe lîngă problemele social-economice, condițiile de viață și muncă etc., interesează, într-o măsură deosebit de importantă, problemele de sănătate, cum ar fi : condițiile igieno-sanitare din cadrul mediului de muncă și trai, antecedentele morbide familiale și ale decedatului, cauza inițială și directă a decesului, asistența medicală înainte de deces (consultații, internări în spitale, tratamente etc.).

Problema numărului eșantioanelor și a mărimii lor este deosebit de importantă, deoarece ea decide, în bună parte, exactitatea rezultatelor obținute din cercetare.

În alegerea eșantionului, a numărului de colectivități din care trebuie să alegem eșantionul, tebuie să ținem seama de condițiile geografice (teritorii muntoase, deluroase, cîmpie, litoral), de condițiile social-economice deosebite în dezvoltarea unei anumite populații și care fac ca unele grupuri de populație să-și aibă specificul lor, care să influențeze fenomenele demografice.

Alegerea eșantioanelor trebuie să țină seama de ocupația de bază a populației (teritorii agricole, industriale) apoi de ramurile de producție în care activează locuitorii din eșantionul ales, populațiile cu natalitate ridicată sau scăzută sau cu migrație puternică sau slabă etc.

Gradul de cultură generală și cultură sanitară, ca și asigurarea bună sau slabă a populației cu asistență medicală constituie, de asemenea, criterii importante de stabilire a numărului eșantionului.

În acest fel, în cazul unei cercetări selective pe o regiune, numărul de eșantioane se ridică la aproximativ 4 — 6 grupe de locuitori, adică 4 — 6 teritorii diferite, circumscripții sanitare, de exemplu. Acest număr de eșantioane nu se poate generaliza în mod uniform la toate regiunile. El depinde de o serie de condiții care determină o variabilitate mai mare sau mai mică a problemelor, o diversitate de caracteristici, care fac ca populația din diferitele teritorii ale regiunii să prezinte un specific al ei propriu, față de populația din alte teritorii ale regiunii.

Aceasta nu înseamnă că ar trebui să creștem prea mult numărul de eșantioane, ci, din contră, uneori putem chiar scădea acest număr în funcție de variabilitate.

De asemenea, atunci cînd luăm eșantioane din diferite teritorii ale regiunii nu avem în față, în mod mecanic, specificul geografic sau istoric etc., ci avem totdeauna în față faptul dacă aceste situații specifice au putut sau nu să ducă la o variabilitate în cadrul problemelor demografice de cercetat. Doar în acest caz vom crește sau nu numărul de eșantioane. Deci numărul eșantionului depinde, în primul rînd, de variabilitatea problemei studiate.

Cele arătate anterior ne duc și la o altă concluzie, și anume : numărul de eșantioane nu se alege conform selecției mecanice sau selecției întîmplătoare, ci prin selecție tipică, ținînd seamă de diferitele probleme specifice din punct de vedere demografic. Vom lua acele raioane și circumscripții care pe baza analizei premergătoare cercetării, pe baza unei documentații serioase, ne atrag atenția asupra situației lor specifice.

În legătură cu aceasta se mai ridică o problemă :

Eșantionul trebuie să prindă ceea ce este „normal", „tipic" sau trebuie să cuprindă diferitele „tipuri" de colectivități ?

Dacă scopul cercetării este de a constata care este fenomenul în ansamblul lui, care este situația caracteristică regiunii sau teritoriului respectiv, indiferent de unele oscilații, variabilități inerente oricărei probleme demografice, în acest caz trebuie să cercetăm acele colectivități care sînt reprezentative, tipice pentru întreaga colectivitate (tipicul trebuie luat ca fiind caracteristic pentru esența materială a fenomenului și nu tipic în sens de „valoare medie", în sensul concepției lui A. Quetelett).

Dacă scopul cercetării este de a cunoaște toate tipurile de populație, indiferent dacă unele din aceste tipuri reprezintă o parte infimă din totalul ei, 1 — 2% eventual, în acest caz desigur că numărul eșantioanelor va fi mai mare.

De exemplu, în cazul unei cercetări care are ca scop studiul cauzelor mortalității, vom ține seamă de ceea ce este tipic, caracteristic majorității populației, unde se constată fenomenul respectiv, mortalitate generală sau infantilă crescută.

O altă problemă care se pune în cadrul cercetării demografice este cea a eșantionului martor, deosebit de important și în cercetările demografice, ca și în cele clinice sau de laborator.

Se pune întrebarea dacă în cercetările demografice este sau nu nevoie de eșantion martor și ce fel de colectivitate trebuie să reprezinte el ?

Se știe că în cercetările de laborator eșantionul martor servește ca unitate de comparație — etalon — pentru celelalte loturi de elemente sau animale de laborator, asupra cărora se efectuează cercetarea. Ființele sau elementele care fac parte din eșantionul martor reprezintă „normalul" cu care se compară restul de eșantioane ce sînt cercetate.

În cazul cercetărilor privind starea de sănătate a populației, în cercetările demografice sau de morbiditate, eșantionul martor este mult mai greu de ales și are și o altă semnificație decît în cercetările de laborator sau chiar și în cele clinice.

În acele cercetări selective, în care cunoaștem de la început la ce rezultate ne putem aștepta, ca de exemplu regiuni sau raioane cu natalitate scăzută sau cu mortalitate generală sau infantilă crescută, ca eșantion martor trebuie să alegem teritorii cu natalitate medie și separat teritorii cu natalitate mare, iar în cazul mortalității, teritorii cu mortalitate scăzută și separat teritorii cu mortalitate crescută.

Numărul de eșantioane va fi, deci, mai mare decît în cazul cercetărilor de laborator și va fi diferit, în sensul că vor fi luate, pe lîngă eșantioanele tipice pentru problema în studiu (adică populații cu natalitate scăzută sau mortalitate generală sau infantilă crescută), și eșantioane martor cu specific deosebit.

Acest eşantion diferit nu trebuie luat numai din teritoriile în care fenomenul de studiat prezintă o valoare medie, dar şi din acele teritorii în care fenomenul este cu totul opus faţă de cel aflat în studiu.

Aşadar eşantionul martor în cercetarea stării de sănătate a populaţiei diferă de cele din cercetările clinice şi de laborator (unde martorul reprezintă „normalul"), atît prin numărul mai mare de eşantioane, cît şi prin diversitatea lor (eşantioane cu intensităţi diferite în ceea ce priveşte problema studiată).

În cazul în care nu avem date documentare în ceea ce priveşte problema pe care o cercetăm, nu vom putea lua nici eşantion martor, neavînd criterii de comparare. Din aceste motive se iau, în general, eşantioane din diferite teritorii cu specific aparte în ceea ce priveşte relieful, clima, dezvoltarea social-istorică, ocupaţia de bază a populaţiei, mediul urban sau rural etc.

Aceste eşantioane în realitate pot fi considerate ca eşantioane martor, unul faţă de celălalt, nemaifiind necesare şi alte eşantioane de comparare.

Revenind la problema numărului de eşantioane, pe care le-am fixat în cadrul unei regiuni, la un număr de aproximativ 4 — 6 circumscripţii (după caracteristicile problemei de cercetat şi după gradul de variabilitate a problemei respective), nu rezultă că cercetarea aprofundată pe un raion poate să se rezume la un singur eşantion (o singură circumscripţie sau comună). Şi în acest caz se aplică cele expuse anterior că va trebui să luăm o serie de eşantioane reprezentative din diferite teritorii, respectiv populaţii ale raionului. Acest număr depinde, de asemenea, de variabilitatea fenomenului din cadrul raionului respectiv.

Dacă fenomenul ar fi identic pe întreaga regiune, ne-am putea rezuma la o singură comună sau circumscripţie. În realitate, fenomenele stării de sănătate sînt de o mare variabilitate şi din aceste motive nu putem să ne mulţumim cu un număr mic de eşantioane, ci de la caz la caz 4 — 6 sau chiar mai multe circumscripţii.

De asemenea, şi cercetarea pe întreaga ţară va trebui să cuprindă un număr de aproximativ 150 de circumscripţii.

Numărul scăzut de teritorii care sînt luate în cercetare se explică prin faptul că nu se ia, în mod mecanic, acelaşi număr de eşantioane din toate regiunile, ci, ţinînd seama de caracteristicile problemei, doar din anumite grupuri de raioane, cu specific diferit unul de celălalt.

În cazul studiului natalităţii şi fecundităţii generale şi conjugale vom lua eşantioane din 1 — 2 regiuni cu natalitatea cea mai scăzută, dintr-o regiune cu natalitatea medie şi dintr-o regiune cu natalitatea crescută, în total deci din 3 — 4 regiuni.

În cazul mortalităţii generale sau, mai ales, în cazul mortalităţii infantile numărul de regiuni va fi mai mare : 3 — 4 regiuni cu mortalitate infantilă ridicată, 2 regiuni cu indici medii şi 1 — 2 regiuni cu indici scăzuţi, deci în total 6 — 8 regiuni.

Desigur că orice tipizare este mecanică, dacă ea nu este efectuată pe baza unei documentări ştiinţifice, iar la sfîrşit verificată prin calcule matematice ale mediei, erorii medii şi a dispersiunii de la medie etc.

După fixarea numărului eşantionului (a numărului de colectivităţi, circumscripţii), trecem la fixarea mărimii eşantionului, adică a numărului de cazuri ce trebuie cercetate de fiecare eşantion.

Mai corect este să se cerceteze întreaga populaţie din eşantionul fixat, în cazul nostru circumscripţie sanitară sau comună, adică să se ia în considerare întreaga colectivitate care formează o unitate complexă social-economică, cultural-sanitară etc.

Nu este just să se ia în cercetare doar cîteva familii sau cîteva gospodării din aceste eşantioane (numite şi segmente), cum s-a procedat de exemplu în ancheta morbidităţii din S.U.A. (1957 — 1959), în care au fost aleşi în studiu doar 6 familii din fiecare segment. În acest caz nu se poate elimina întîmplătorul, eşantioanele nefiind caracteristice pentru întreaga colectivitate, după cum nici aceste 6 familii nu sînt caracteristice pentru eşantionul ales.

C. METODE DE APRECIERE A OMOGENITĂŢII EŞANTIONULUI ŞI A GRADULUI DE VARIABILITATE A FENOMENELOR DEMOGRAFICE

Să luăm un exemplu referitor la letalitatea printr-o anumită boală (X), letalitatea avînd formula următoare :

$$\left(\frac{\text{numărul deceselor prin boala X} \cdot 100}{\text{numărul total de decese}}\right).$$

Să presupunem că eşantionul provine din 50 de circumscripţii cu un număr total de 2 600 de decese generale pe un an, circumscripţiile fiind alese pe bază de studiu prealabil, considerate ca fiind resprezentative pe întreaga ţară. Din cele 2 600 de decese generale, 104 au fost decese prin boala X, adică o letalitate de 4,0% $\left(\frac{104 \cdot 100}{2\,600} = 4{,}0\%\right)$.

Pentru a putea aprecia omogenitatea eşantionului şi gradul de variabilitate şi deci dacă el este reprezentativ pentru întreaga ţară (adică pentru a putea afirma că letalitatea pe ţară prin boala X este de 4,0%), vom calcula deviaţia-standard, coeficientul de variaţie, eroarea medie a mediei aritmetice şi procentul erorii medii :

$$\sigma = \pm \sqrt{\frac{P \cdot (100 - P)}{N}} = \pm \sqrt{\frac{4 \cdot 96}{2\,600}} = \pm\ 0{,}38\,;$$

$$c\,V = \pm \frac{\sigma \cdot 100}{Mp} = \pm \frac{0{,}38 \cdot 100}{4} = \pm\ 9{,}5\%.$$

$$e\,M = \pm \frac{\sigma}{\sqrt{N}} = \pm \frac{0{,}38}{\sqrt{2\,600}} = \pm \frac{0{,}38}{50{,}99} = \pm\ 0{,}007\,;$$

$$e\,M\% = \pm \frac{e\,M \cdot 100}{Mp} = \pm \frac{0{,}007 \cdot 100}{4{,}0} = \pm\ 0{,}18\%.$$

Se poate constata din indicii mici ai deviaţiei-tip ($\sigma = \pm\ 0{,}38$) şi din coeficientul de variaţie ($cV = 9{,}5\%$) că eşantionul ales este omogen, că gradul de dispersiune nu este mare (sub 10%). De asemenea, din indicii

erorii medii ($eM = \pm 0{,}007$) şi procentul erorii medii ($eM\% = \pm 0{,}18\%$) se constată că letalitatea de 4,0% prin boala X este reprezentativă pentru letalitatea pe toată ţara, indicele pe ţară putînd fi cuprins cu o mare probabilitate (99,73%) între acest indice de 4,0% $\pm 3 . 0{,}007$ (deci de trei ori eroarea medie). Aceasta înseamnă că indicele de letalitate pe ţară va putea fi cuprins între 3,98% şi 4,02%, ceea ce denotă că eşantionul a fost bine ales, avînd o medie reprezentativă pentru întreaga ţară.

D. METODE DE APRECIERE A CERTITUDINII DIFERENŢEI DINTRE FRECVENŢELE RELATIVE ÎN DEMOGRAFIE

În cazul cercetărilor parţiale cu privire la mişcarea naturală a populaţiei este necesar, de multe ori, să se compare doi indici de natalitate sau mortalitate generală sau infantilă etc., dacă diferenţa dintre ei este sau nu reprezentativă. Acest fapt este necesar nu numai comparativ între două teritorii, ci mai ales în comparaţie dinamică : în ce măsură un indice mai mic de mortalitate infantilă, de exemplu faţă de un an anterior, denotă o îmbunătăţire reală sau dacă această scădere a indicelui este doar întîmplătoare, din cauza eşantioanelor diferit alese.

Să luăm exemplul mortalităţii infantile de 6,3% dintr-un an oarecare şi indicele de 6,1% din anul următor, calculaţi pe baza a 90 de eşantioane cu un total de 240 de decese sub 1 an pe fiecare an studiat. Să vedem în ce măsură scăderea este reală sau, din contră, neesenţială, întîmplătoare, din cauza eşantioanelor nereprezentative ce au fost alese în cei doi ani consecutivi.

Se calculează deviaţia-tip la cele două medii, conform formulei cunoscute la calculul dispersiunii mărimilor relative, notînd deviaţia-tip din primul an cu σ_1 şi cea din anul următor cu σ_2.

$$\sigma_1 = \pm \sqrt{\frac{P_1 \cdot (100 - P_1)}{N_1}} = \pm \sqrt{\frac{6{,}3 \cdot 93{,}7}{240}} = \pm \sqrt{2{,}46} = \pm 1{,}57 ;$$

$$\sigma_2 = \pm \sqrt{\frac{P_2 \cdot (100 - P_2)}{N_2}} = \pm \sqrt{\frac{6{,}1 \cdot 93{,}9}{240}} = \pm \sqrt{2{,}39} = \pm 1{,}54.$$

Pentru a calcula dacă diferenţele celor doi indici (6,3% şi 6,1%) sînt sau nu certe, se calculează erorile medii :

$$e\,M_1 = \pm \frac{\sigma 1}{\sqrt{N_1}} = \pm \frac{1{,}57}{\sqrt{240}} = \pm \frac{1{,}57}{15{,}49} = \pm 0{,}10.$$

$$e\,M_2 = \pm \frac{\sigma_2}{\sqrt{N_2}} = \pm \frac{1{,}54}{\sqrt{240}} = \pm \frac{1{,}54}{15{,}49} = \pm 0{,}10$$

Calculăm, în continuare, diferenţa între cele două medii ale mortalităţii infantile : 6,3% — 6,1% = 0,2%.

Calculăm, de asemenea, şi eroarea diferenţei după formula cunoscută în capitolul I.

$$e\,D = \pm \sqrt{e\,M_1^2 + e\,M_2^2} = \pm \sqrt{0{,}10^2 + 0{,}10^2} = \pm 0{,}14.$$

Se constată că diferenței dintre cele două medii ale mortalității infantile (0,2%) îi corespunde o eroare medie a diferențelor de $\pm$ 0,14, iar triplul acestuia este de 0,42, mai mare decît diferența dintre mediile mortalității infantile.

Se poate, deci, afirma cu o probabilitate de 99,73% că scăderea mortalității infantile este doar întîmplătoare și nu reală.

Să luăm un alt exemplu pentru calculul certitudinii diferenței a doi indici de mortalitate infantilă pe un număr de 120 de eșantioane cu un număr total de 300 de decese sub 1 an, anual. În primul an mortalitatea infantilă a fost de 6,5%, iar în anul următor de 6,0%.

Calculăm deviațiile-standard și erorile medii ale mediei aritmetice după formulele anterioare:

$$\sigma_1 = \pm \sqrt{\frac{6{,}5 \cdot 93{,}5}{300}} = \pm \sqrt{2{,}03} = \pm 1{,}42;$$

$$\sigma_2 = \pm \sqrt{\frac{6 \cdot 94}{300}} = \pm \sqrt{1{,}88} = \pm 1{,}37;$$

$$e\,M_1 = \pm \frac{1{,}42}{\sqrt{300}} = \pm \frac{1{,}42}{17{,}32} = \pm 0{,}082;$$

$$e\,M_2 = \pm \frac{1{,}37}{\sqrt{300}} = \pm \frac{1{,}37}{17{,}32} = \pm 0{,}079.$$

Diferența dintre cei doi indici de mortalitate infantilă este de 6,5% — 6,0% = 0,5%.

Calculăm eroarea diferenței:

$$e\,D = \pm \sqrt{0{,}082^2 + 0{,}079^2} = \pm \sqrt{0{,}013} = \pm 0{,}11.$$

Constatăm că triplul erorii diferenței $3 \cdot (\pm 0{,}11) = \pm 0{,}33$ este mai mic decît diferența între cei doi indici de mortalitate infantilă (0,5%), ceea ce ne permite să concludem cu o probabilitate de 99,73%, deci cu o eroare doar de 0,27%, că mortalitatea infantilă a scăzut în mod cert și nu întîmplător.

E. METODE DE CALCUL A EȘANTIONULUI REPREZENTATIV ÎN DEMOGRAFIE

În exemplele anterioare am ales numărul de eșantioane pe baza unui studiu prealabil, ținînd seama de o serie de condiții pentru ca eșantionul să fie cît mai reprezentativ pentru problema și regiunea studiată.

Să luăm, de exemplu, o mortalitate infantilă de 7,0% și de 5,0% din anul următor, studiate pe baza unui număr anual de 2 000 decese sub 1 an și să calculăm dacă acest eșantion este suficient de reprezentativ pentru a trage concluzia generală că mortalitatea infantilă este într-adevăr în scădere sau dacă scăderea de la 7,0 la 5,0% este doar întîmplătoare, neesențială.

Se utilizează și în acest caz calculul deviației-tip a mărimilor relative și eroarea diferenței.

$$\sigma_1 = \pm\sqrt{\frac{P_1 \cdot (100 - P_1)}{N_1}} \qquad \sigma_2 = \pm\sqrt{\frac{P_2 \cdot (100 - P_2)}{N_2}}.$$

Diferența dispersiunii va fi ca și în cazul erorii medii a diferenței :

$$\sigma D = \pm\sqrt{\sigma_1^2 + \sigma_2^2} = \pm\sqrt{\frac{P_1 \cdot (100 - P_1)}{N_1} + \frac{P_2 \cdot (100 - P_2)}{N_2}}.$$

Deoarece $N_1 = N_2$, formula va fi următoarea :

$$\sigma D = \pm\sqrt{\frac{P_1 \cdot (100 - P_1) + P_2 \cdot (100 - P_2)}{N}}$$

Dacă vom lua datele exemplului nostru și le vom situa în formula anterioară vom obține :

$$\sigma D = \pm\sqrt{\frac{7{,}0 \cdot 93{,}0 + 5 \cdot 95}{2\,000}} = \pm\sqrt{\frac{651{,}0 + 475}{2\,000}} = \pm\sqrt{0{,}50} = \pm\, 0{,}75.$$

Triplul erorii medii a diferenței va fi : $3. (\pm\, 0{,}75) = 2{,}25$.

Această diferență este mai mare decît diferența dintre cei doi indici ai mortalității infantile : 7,0%—5,0% = 2,0%.

Pentru ca această diferență de 2,0% să fie certă și nu întîmplătoare ar trebui să creștem eșantionul global (N) pentru ca triplul erorii medii a diferenței să nu fie mai mare, ci mai mic decît diferența dintre cei doi indici ai mortalității infantile. Aceasta înseamnă că eroarea medie a diferenței nu poate crește peste 0,66, care înmulțită cu 3 dă 1,98, deci sub valoarea lui 2,0%.

Numărul eșantionului necesar, adică reprezentativ, se calculează pe baza formulei anterioare a diferenței dispersiunii (σD), de unde calculăm pe N :

$$N = \frac{P_1(100 - P_1) + P_2(100 - P_2)}{\sigma D^2} = \frac{7{,}0 \cdot 93{,}0 + 5 \cdot 95}{0{,}66^2} = \frac{1126{,}0}{0{,}44} = 2559.$$

Rezultă că numărul eșantionului, cele 2 000 de decese anuale la copiii sub 1 an, nu este suficient.

Dacă ne mulțumim cu o probabilitate de 95,45%, numărul de 2 000 de decese sub 1 an ar fi fost chiar peste necesar, eșantionul reprezentativ (N) fiind, în acest caz, de 1 126.

Astfel σD ar fi 1,0, care înmulțit cu 2 ar fi egal cu diferența dintre cei doi indici (7,0% — 5,0% = 2,0%). Formula eșantionului reprezentativ în acest caz este următoarea :

$$N = \frac{7{,}0 \cdot 93{,}0 + 5{,}0 \cdot 95{,}0}{1^2} = \frac{1\,126{,}0}{1} = 1\,126.$$

Deci eşantionul reprezentativ trebuie să fie cu atît mai mare, cu cît dispersiunile celor două eşantioane sînt mai mari şi, invers, cu cît dispersiunile a două eşantioane sînt mai mici cu atît şi eşantionul reprezentativ va putea fi mai mic.

F. METODE DE CORECTARE A INDICILOR DEMOGRAFICI

În calculul diferiţilor indici demografici, ca şi de morbiditate, de altfel, atunci cînd se studiază comparativ aceste fenomene pe diferite regiuni, intervine o eroare datorită neomogenităţii populaţiei, din punctul de vedere al structurii sale pe sexe, grupe de vîrstă etc.

Astfel în cazul natalităţii sau al mortalităţii generale, indicii calculaţi pot fi diferiţi, datorită structurii deosebite pe grupe de vîrstă a populaţiei. În cazul natalităţii, indicii calculaţi vor depinde şi de distribuţia pe sexe şi starea civilă a populaţiei.

Să luăm un exemplu concret : mortalitatea generală din regiunea Banat şi regiunea Argeş. Indicii mortalităţii generale pentru aceste două regiuni ne arată un nivel mai ridicat în prima regiune, faţă de a doua, şi anume 12,6‰ în regiunea Banat şi 10,4‰ în regiunea Argeş, pe anul 1957. De altfel, această diferenţă de nivel nu este accidentală, ci ea se menţine ca atare de mai mulţi ani.

Se poate oare trage concluzia că mortalitatea generală, mai ridicată în regiunea Banat, indică o stare de sănătate mai nesatisfăcătoare faţă de regiunea Argeş? Desigur că nu, căci indicele mai crescut se datoreşte şi structurii populaţiei pe grupe de vîrstă, existînd o proporţie mult mai mică de populaţie tînără şi o proporţie mult mai mare de bătrîni decît în cazul regiunii Argeş. Or, mortalitatea este mult mai mare la bătrîni, aceasta influenţînd, prin ponderea ei, mortalitatea generală globală. Natalitatea scăzută zeci de ani în regiunea Banat a dus la scăderea grupelor tinere de vîrstă şi creşterea celor bătrîne. Aceasta rezultă concret din următoarele date : populaţia de peste 50 de ani la recensămîntul din 1930 în Banat forma 20,7% din total, în timp ce la recensămîntul din 1956, populaţia de peste 50 de ani a regiunii Banat constituia 27,6% din total, deci o creştere a populaţiei de bătrîni, ceea ce a influenţat creşterea mortalităţii generale.

Să luăm un alt exemplu concret : mortalitatea generală din regiunea Maramureş, comparativ cu cea din regiunea Ploieşti, indicele fiind mai crescut în prima regiune, şi anume 10,8‰ faţă de 9,2‰ în regiunea Ploieşti în anul 1957. Şi în cazul regiunii Maramureş indicele acesta ridicat nu este accidental, ci se menţine de mai mulţi ani.

Se poate oare afirma că mortalitatea generală mai mare în regiunea Maramureş se datoreşte structurii pe grupe de vîrstă? Este greu de răspuns în mod categoric, deoarece natalitatea nu are nivelul scăzut, cum este în cazul regiunii Banat. S-ar părea că în acest caz mortalitatea generală mai mică în regiunea Ploieşti s-ar datori unei situaţii mai favorabile a stării de sănătate. Rămîne, totuşi, ca o afirmaţie slab dovedită.

Deci, indicii globali de natalitate şi mortalitate generală nu reflectă, în general, realitatea (intensitatea natalităţii, intensitatea mortalităţii) în cazul comparării a două sau mai multe regiuni cu o structură diferită a populaţiei pe grupe de vîrstă.

Din aceste motive, în calcularea natalității, mortalității sau morbidității, se utilizează, pe lîngă indicii globali, și indicii corectați, pe baza metodelor de standardizare a populației îndeosebi.

Standardizarea populației se face luînd ca bază populația unei anumite țări sau regiuni, unde structura pe sexe sau grupe de vîrstă nu s-a modificat pe o perioadă îndelungată sau la care modificările n-au fost prea mari. Dar aceasta nu este o regulă, căci se poate utiliza ca standard populația oricărei țări sau oricărei regiuni, indiferent de distribuția populației sale.

Pentru studiul evolutiv pe o perioadă îndelungată, peste 20 de ani de exemplu, este necesar să utilizăm, fie populația standard a unei țări sau regiuni a cărei structură nu s-a modificat prea mult în această perioadă, fie că utilizăm populația standard dintr-o țară sau regiune dintr-un anumit an, pentru calculul indicilor din toți anii perioadei respective.

Este deci greșit a utiliza, așa cum se obișnuiește, de multe ori, ca populație standard, populația unor țări în diferiți ani, căci de la o perioadă la alta compoziția populației se schimbă și indicii comparativi nu mai reflectă realitatea.

Astfel, de exemplu, pentru studiul mortalității corectate pe regiuni se va lua ca populație standard, fie cea din anul 1956, fie cea din anul 1930 (ani de recensămînt) pentru a putea avea o bază de comparație unică.

Desigur că pentru populația standard nu este necesar să luăm neapărat ani de recensămînt, ci orice an în care cunoașterea structurii populației este posibilă, în cazul în care mișcarea naturală și mecanică a populației se urmărește în mod continuu pe această structură.

Este mai bine însă a se utiliza populația la recensăminte, datele fiind mult mai exacte.

a) **Metode de corectare a indicelui de natalitate.** Scăderea continuă a natalității și menținerea ei la un nivel scăzut timp de decenii duce la schimbarea structurii pe grupe de vîrstă a populației, în sensul scăderii grupelor tinere de vîrstă și creșterii grupelor vîrstnice. Această structură modificată poate duce, la rîndul ei, la scăderea în continuare a natalității.

Deoarece între vîrsta de 20 și 30 de ani a mamelor au loc cele mai multe nașteri, se înțelege că populația cu un procent mai mare de femei de această vîrstă va putea avea o natalitate mai mare, ținînd seama, bineînțeles, de factorii social-economici, și alți factori care determină nivelul și dinamica natalității.

În orice caz comparația indicilor de natalitate între două țări sau regiuni nu exprimă în mod real fertilitatea populației, în cazul diferențelor mari în structura ei pe sexe și pe grupe de vîrstă.

Corectarea natalității se face prin calcularea fertilității generale și conjugale, în care caz corecția ține seamă de sexul și starea civilă a populației.

Formulele de calcul ale fertilității generale și conjugale sînt :

$$F\,gl = \frac{\text{Numărul născuților vii} \cdot 1\,000}{\text{numărul femeilor în vîrstă de } 15 - 49 \text{ de ani}}$$

$$Fc = \frac{\text{Numărul născuților vii legitimi} \cdot 1\,000}{\text{Numărul femeilor căsătorite}}$$

Fertilitatea generală şi conjugală, pe ani de vîrstă, introduce şi corecţia pentru grupele de vîrstă ale femeilor, deci va măsura mult mai exact intensitatea specifică a natalităţii populaţiei decît în cazul indicelui global al fertilităţii.

$$Fc\ (20-24\ \text{de ani}) = \frac{\text{Numărul născuţilor vii legitimi ai mamelor în vîrstă de 20 — 24 de ani} \cdot 1\,000}{\text{Numărul femeilor în vîrstă de 20 — 24 de ani, căsătorite}}$$

Fertilitatea generală sau conjugală se mai poate calcula după vîrsta mamei şi după ordinea de naştere a copilului, adică dacă este primul copil născut viu sau al doilea etc. Se calculează la numărul de femei între 15 şi 49 de ani sau în cazul fertilităţii conjugale la numărul de femei căsătorite la grupa de vîrstă de 15—49 ani.

Formulele de calcul sînt următoarele :

$$F\,gl = \frac{\text{Numărul primilor născuţi vii ai femeilor în vîrstă de 15—49 de ani} \cdot 1\,000}{\text{Numărul femeilor în vîrstă de 15—49 de ani}}$$

$$Fc = \frac{\text{Numărul primilor născuţi vii ai femeilor în vîrstă de 15—49 de ani, căsătorite} \cdot 1\,000}{\text{Numărul femeilor în vîrstă de 15—49 de ani, căsătorite}}$$

La fel se calculează şi pentru al doilea născut viu, al treilea etc.

Dăm, în continuare, în tabelul XX, indicii de fertilitate generală şi ordinea de născuţi vii în cîteva ţări.

Tabelul XX

Indicele de fertilitate generală [1], după ordinea născuţilor vii [2]

Ordinea născuţilor vii	Guatemala (1950)	S.U.A. (1953)	Austria (1953)	Franţa (1953)	Olanda (1953)	Portugalia (1954)	Elveţia (1953)	Iugoslavia (1954)
Total	170,1	85,7	42,7	65,0	76,8	65,6	59,0	92,3
1	45,1	24,3	16,4	19,9	22,1	19,4	22,2	28,9
2	33,4	23,6	12,9	17,4	18,7	13,5	16,0	23,7
3	26,9	15,8	6,6	11,8	12,7	9,7	9,7	15,1
4	19,2	8,6	3,2	6,9	8,3	6,9	5,1	9,1
5	13,7	4,5	1,6	3,8	5,4	5,1	2,7	5,5
6	9,8	2,5	0,8	2,0	3,5	3,5	1,5	3,5
7	6,3	1,5	0,4	1,2	2,1	2,6	0,8	2,4
8	4,5	0,9	0,3	0,7	1,4	1,8	0,5	1,5
9	2,8	0,6	0,2	0,4	1,0	1,2	0,3	1,0
10 şi peste	3,6	1,0	0,3	0,7	1,5	1,8	0,4	1,4
Nespecificat	5,0	2,4	0,0	0,2	—	0,0	—	0,1

[1] Indicele de fertilitate generală s-a calculat la populaţia feminină în vîrstă de 10—49 de ani.
[2] Annuaire Demographique, New York, 1955, tabel 28, p. 539.

b) **Metode de corectare a indicelui de mortalitate generală.** În cazul mortalităţii generale se utilizează indicele de mortalitate corectată. El se calculează după diferite metode.

Metoda cea mai simplă este corectarea mortalităţii prin raportarea ei la o populaţie aleasă ca standard, în locul populaţiei teritoriului respectiv.

Se calculează mai întîi procentul de locuitori pe fiecare an sau pe grupe de vîrstă din totalul populaţiei. Se ia apoi un număr de un milion de locuitori, la care se aplică aceste procente, şi se află, în acest fel, cifrele absolute pe aceeaşi grupă de vîrstă.

Obţinem, astfel, populaţia standard de un milion de locuitori (tabelul XXI), pe care o vom utiliza la calcularea mortalităţii corectate, în exemplul nostru într-un raion oarecare B.

Tabelul XXI

Stabilirea populaţiei standard din regiunea A. (date arbitrare)

Anul de vîrstă	Proporţia la sută din totalul populaţiei	Populaţia standard la un milion de locuitori	Anul de vîrstă	Proporţia la sută din totalul populaţiei	Populaţia standard la un milion de locuitori	Anul de vîrstă	Proporţia la sută din totalul populaţiei	Populaţia standard la un milion de locuitori
Total	100,00	1 000 000						
Sub 1	2,20	22 000						
1	2,18	21 800	34	1,35	13 500	67	0,56	5 600
2	2,12	21 200	35	1,38	13 800	68	0,50	5 000
3	2,09	20 900	36	1,12	11 200	69	0,42	4 200
4	2,04	20 400	37	0,82	8 200	70	0,57	5 700
5	2,03	20 300	38	0,88	8 800	71	0,35	3 500
6	2,02	20 200	39	1,07	10 700	72	0,40	4 000
7	1,98	19 800	40	1,17	11 700	73	0,36	3 600
8	1,61	16 100	41	1,16	11 600	74	0,30	3 000
9	1,60	16 100	42	1,31	13 100	75	0,32	3 200
10	1,56	15 600	43	1,32	13 200	76	0,24	2 400
11	1,57	15 700	44	1 29	12 900	77	0,17	1 700
12	1,55	15 500	45	1,25	12 500	78	0,18	1 800
13	1,54	15 400	46	1,22	12 200	79	0,12	1 200
14	1,52	15 200	47	1,18	11 800	80	0,15	1 500
15	1,51	15 100	48	1,16	11 600	81	0,08	800
16	1,81	18 100	49	1,14	11 400	82	0,08	800
17	1,84	18 400	50	1,19	11 900	83	0,06	600
18	1,82	18 200	51	1,07	10 700	84	0,05	500
19	1,79	17 900	52	1,11	11 100	85	0,05	500
20	1,78	17 800	53	1,03	10 300	86	0,04	400
21	1,75	17 500	54	1,01	10 100	87	0,02	200
22	1,82	18 200	55	1,11	11 100	88	0,02	200
23	1,86	18 600	56	1,04	10 400	89	0,01	100
24	1,88	18 800	57	0,88	8 800	90	0,02	200
25	1,92	19 200	58	0,89	8 900	91	0,006	60
26	1,85	18 500	59	0,74	7 400	92	0,006	60
27	1,77	17 700	60	0,74	7 400	93	0,004	40
28	1,76	17 600	61	0,63	6 300	94	0,003	30
29	1,79	17 900	62	0,71	7 100	95	0,004	40
30	2,07	20 700	63	0,66	6 600	96	0,003	30
31	1,56	15 600	64	0,60	6 000	97	0,001	10
32	1,63	16 300	65	0,67	6 700	98	0,001	10
33	1,59	15 900	66	0,59	5 900	99	0,0007	7
						100 şi peste	0,002	20

Pentru calculul mortalităţii corectate utilizăm tabelul XXII.

În prima coloană se trec anii de vîrstă, în coloana următoare se trece populaţia standard, calculată la un milion de locuitori din regiunea A.

Tabelul XXII

Mortalitatea corectată în raionul B. (date arbitrare)

Ani de vîrstă	Populaţia standard din reg. A. (la un milion de locuitori)	Mortalitatea pe ani de vîrstă în raionul B., înmulţită cu populaţia standard	Ani de vîrstă	Populaţia standard din reg. A. (la un milion de locuitori)	Mortalitatea pe ani de vîrstă în raionul B., înmulţită cu populaţia standard
Total	1 000 000	11 432 066			
Sub 1	22 000	2 460 648	51	10 700	90 600
1	21 800	244 640	52	11 100	107 115
2	21 200	81 204	53	10 300	100 116
3	20 900	60 192	54	10 100	128 520
4	20 400	33 864	55	11 100	118 584
5	20 300	28 470	56	10 400	171 808
6	20 200	20 736	57	8 800	127 680
7	19 800	21 360	58	8 900	125 955
8	16 100	21 280	59	7 400	126 688
9	16 000	19 964	60	7 400	153 502
10	15 600	10 350	61	6 300	200 151
11	15 700	12 495	62	7 100	189 996
12	15 500	9 610	63	6 600	235 422
13	15 400	10 890	64	6 000	195 780
14	15 200	11 622	65	6 700	199 057
15	15 100	18 684	66	5 900	231 339
16	18 100	16 926	67	5 600	218 176
17	18 400	15 624	68	5 000	244 200
18	18 200	42 262	69	4 200	237 048
19	17 900	27 808	70	5 700	271 947
20	17 800	30 264	71	3 500	231 630
21	17 500	43 690	72	4 000	282 160
22	18 200	30 770	73	3 600	260 424
23	18 600	50 406	74	3 000	307 440
24	18 800	27 150	75	3 200	248 448
25	19 200	49 104	76	2 400	284 840
26	18 500	29 535	77	1 700	219 606
27	17 700	40 716	78	1 800	215 694
28	17 600	29 356	79	1 200	165 312
29	17 900	32 724	80	1 500	199 185
30	20 700	26 596	81	800	146 744
31	15 600	24 843	82	800	126 904
32	16 300	36 186	83	600	122 196
33	15 900	44 902	84	500	87 750
34	13 500	45 630	85	500	82 170
35	13 800	46 176	86	400	94 204
36	11 200	40 176	87	200	52 632
37	8 200	19 040	88	200	47 058
38	8 800	24 768	89	100	38 272
39	10 700	25 755	90	200	63 736
40	11 700	47 619	91	60	13 830
41	11 600	36 540	92	60	10 385
42	13 100	63 206	93	40	8 438
43	13 200	64 014	94	30	4 286
44	12 900	61 920	95	40	7 895
45	12 500	76 041	96	30	8 000
46	12 200	79 375	97	10	3 333
47	11 800	105 256	98	10	3 333
48	11 600	81 918	99		
49	11 400	111 036	100	20	
50	11 900	29 736			

În coloana a 3-a se trece produsul dintre mortalitatea pe ani de vîrstă din raionul B. şi populaţia standard din coloana a 2-a.

Se face suma coloanei a 3-a şi rezultatul se împarte la un milion pe baza formulei următoare :

$$Mc = \frac{\Sigma Mx \cdot Sx}{1\,000\,000}$$

Mc = mortalitatea corectată ;
Σ = suma ;
Mx = mortalitatea pe ani de vîrstă ;
Sx = populaţia standard pe ani de vîrstă.
Mortalitatea corectată în raionul *B.* va fi deci :

$$Mc = \frac{11\,432\,666}{1\,000\,000} = 11{,}4\,\%$$

Mortalitatea corectată redă mult mai just intensitatea mortalităţii, deci exprimă mult mai bine starea de sănătate a populaţiei, iar comparaţia între regiuni devine posibilă, nu ca în cazul mortalităţii generale.

Dăm în tabelul XXIII mortalitatea corectată pe regiuni [1], paralel cu mortalitatea generală. Ca populaţie standard am utilizat populaţia ţării la recensămîntul din 1956.

Tabelul XXIII

Mortalitatea corectată pe regiuni (1956), comparativ cu mortalitatea generală (proporţii la 1 000 de locuitori)

Regiunea	Mortalitatea		Regiunea	Mortalitatea	
	Generală	Corectată		Generală	Corectată
Argeş	10,4	10,6	Hunedoara	11,4	11,1
Bacău	9,0	10,2	Iaşi	10,0	10,3
Banat	12,9	9,8	Maramureş	11,2	11,4
Braşov	9,4	8,9	Mureş-Autonomă Maghiară	10,4	9,3
Bucureşti	9,3	9,6	Oltenia	10,6	10,5
Cluj	10,3	10,3	Ploieşti	8,6	8,8
Crişana	11,5	10,9	Suceava	9,5	10,1
Dobrogea [2]	9,8	9,9	Oraşul Bucureşti	6,9	7,6
Galaţi	9,5	9,2			

Se poate constata, din acest tabel, că mortalitatea mai mare în regiunea Banat (12,9‰ în anul 1956) se datoreşte în special structurii populaţiei pe grupe de vîrstă, ca dovadă că mortalitatea corectată este printre cele mai mici din ţară (9,8‰), în timp ce regiunea Argeş are o mortalitate corectată mai mare (10,6‰), deşi mortalitatea generală este mai mică decît a regiunii Banat.

De asemenea, se poate vedea că mortalitatea generală este mult mai scăzută în regiunea Ploieşti decît în regiunea Maramureş, ca dovadă că şi după corectare, regiunea Ploieşti are un indice mult mai scăzut (8,8‰ la Ploieşti, faţă de 11,4‰ la Maramureş).

[1] Teritoriul din anul 1956.
[2] Inclusiv oraşul Constanţa.

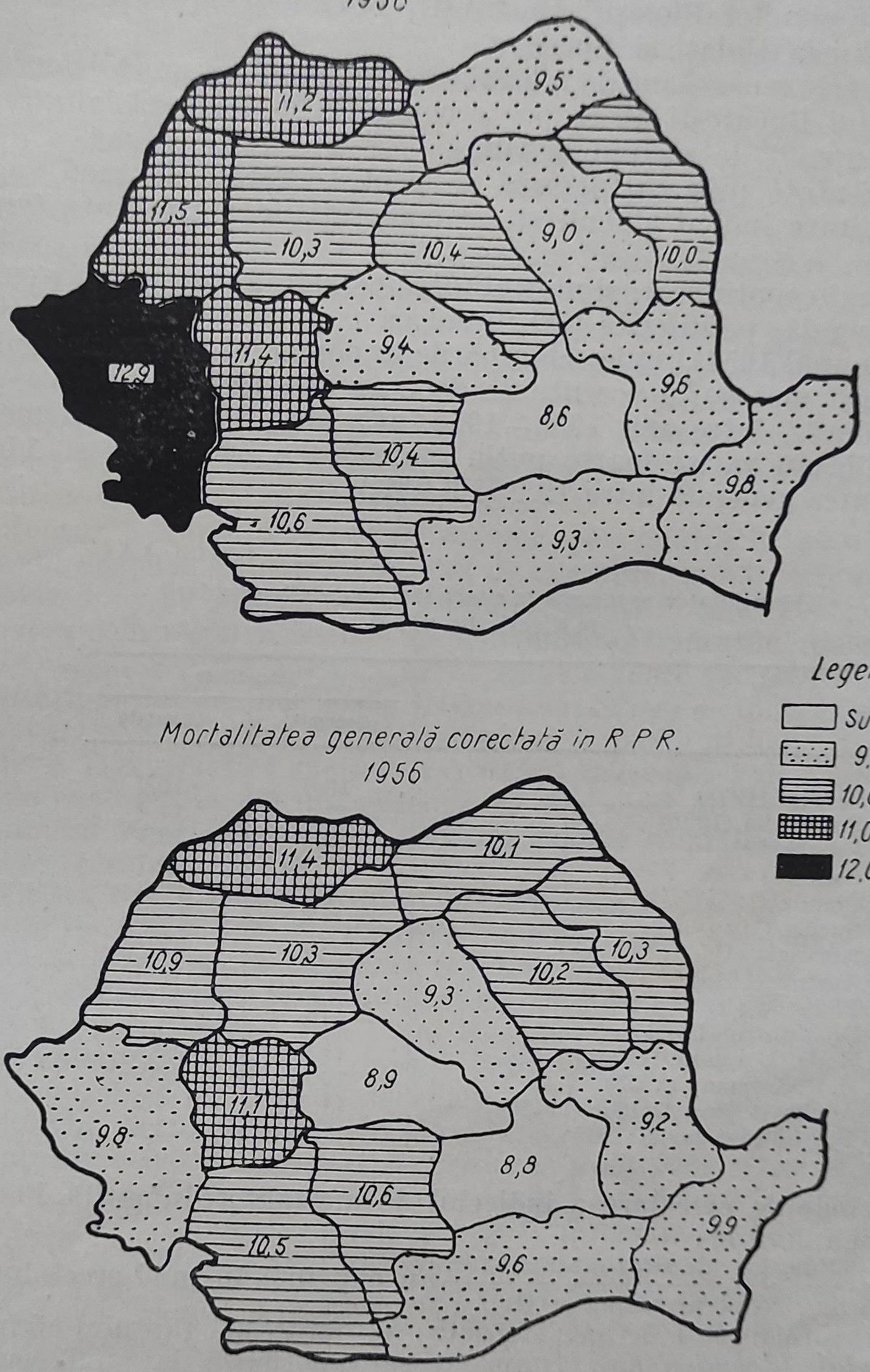

Fig. 9. — **Mortalitatea generală și corectată în R.P.R.**

Regiunea cu cea mai scăzută mortalitate corectată după Ploieşti este Braşov (8,9‰), deşi indicele de mortalitate generală este mai mare decît cel al regiunilor Ploieşti, Bucureşti, Bacău şi aproape la acelaşi nivel cu regiunea Galaţi şi Suceava.

Indicele cel mai scăzut de mortalitate generală şi mortalitate corectată a oraşului Bucureşti nu se poate compara însă cu celelalte regiuni, fiind vorba numai de mediul urban.

Reprezentate prin cartogramă se poate constata în mod sugestiv diferenţele dintre indicii globali şi cei corectaţi de mortalitate (graficul din fig. 9).

Redăm, în continuare, şi mortalitatea corectată, comparativ cu mortalitatea generală, pe diferite ţări, luînd ca populaţie standard pe cea a Elveţiei (din anul 1954), unde schimbările în domeniul fenomenelor demografice au fost mai puţin accentuate decît în alte ţări.

Într-adevăr, între anii 1930—1955, atît natalitatea, cît şi mortalitatea generală au variat foarte puţin (natalitatea între 20,0 şi 17,0‰, iar mortalitatea generală între 12,2 şi 9,9‰).

Tabelul XXIV

Mortalitatea corectată în cîteva ţări (populaţia standard în Elveţia din 1954)

Ţările	Mortalitate	
	Generală	Corectată
Elveţia (1955)	10,0	10,0
Portugalia (1956)	12,1	14,4
Franţa (1956)	12,5	10,5
Belgia (1955)	12,3	10,7
Olanda (1955)	7,6	8,5
Suedia (1955)	9,5	8,6
Norvegia (1955)	8,5	7,7
Austria (1956)	12,4	11,2
Finlanda (1956)	9,0	11,9
Danemarca (1955)	8,7	8,8
R. D. Germană (1955)	11,9	10,4
R. F. Germană (1955)	10,8	10,9
Anglia şi Ţara Galilor (1955)	11,7	10,2
R.P.F. Iugoslavia (1955)	11,3	13,6

c) **Metode de corectare a indicelui de mortalitate infantilă.** Formula mortalităţii infantile :

$$Mi = \frac{\text{Numărul deceselor sub 1 an} \cdot 100}{\text{Numărul născuţilor vii}}$$

nu exprimă în mod precis intensitatea mortalităţii sub 1 an. Aceasta se datoreşte faptului că în statistica oficială decesele sub 1 an nu se raportează la anul naşterii, ci la anul calendaristic al decesului. Deci, dacă un copil s-a născut în februarie 1962, şi a murit în ianuarie 1963, el intră în categoria deceselor sub 1 an din anul 1963. În acelaşi timp, însă, născuţii vii, la care se raportează cei decedaţi sub 1 an din anul 1963, se referă numai la cei născuţi în anul 1963, nu şi la cei din anul 1962.

De aici urmează că numărătorul şi numitorul sînt două entităţi diferite, care nu pot fi raportate una la cealaltă decît cu riscul unei erori.

Calculînd însă mortalitatea infantilă an de an, prin aceeaşi formulă, eroarea de calcul îşi pierde importanţa în mod practic.

În regiunile cu mortalitate infantilă mică, unde peste jumătate din decesele sub 1 an au loc în prima lună a vieţii, această eroare devine insignifiantă, calculul redînd, în mod destul de precis, intensitatea mortalităţii sub 1 an.

În regiunile cu mortalitate infantilă mai mare, unde o mare parte a deceselor survin după prima lună a vieţii şi chiar după luna a VI-a, formula duce la o eroare mult mai mare.

Eroarea devine şi mai accentuată în cazul variaţiilor mai mari ale natalităţii de la un an la altul, deoarece raportăm un număr important din decesele sub 1 an la copii născuţi în anul anterior la născuţii vii din anul următor, la care se calculează indicele şi cînd s-ar putea întîmpla ca natalitatea să fie mult mai scăzută sau mult mai crescută, ţinînd seama de oscilaţiile de la un an la altul al indicilor demografici.

S-au căutat multe măsurători ale intensităţii mortalităţii infantile, în vederea corectării erorilor de calcul.

Metoda cea mai justă ar fi, desigur, prelucrarea deceselor sub 1 an după anul naşterii, pentru a raporta, de exemplu, la născuţii vii din anul 1963 doar decesele sub 1 an ale copiilor născuţi în anul 1963, iar ceilalţi să fie recalculaţi (retrospectiv) la mortalitatea infantilă din anul 1962.

Deoarece această metodă ar necesita o mare întîrziere a calculării mortalităţii infantile, nu poate fi considerată ca metodă practică, deşi ca cercetare retrospectivă pe mai mulţi ani se poate utiliza.

Este mai practică metoda raportării deceselor sub 1 an din anul 1963, de exemplu la 2/3 din născuţii vii din acest an şi 1/3 din născuţii vii ai anului 1962 . Metoda nu necesită calcule complicate şi nu duce la întîrziere în calcularea indicelui.

Există totuşi o deficienţă şi în utilizarea acestei metode, deoarece proporţia de 2/3 şi 1/3 din născuţii vii este aleasă arbitrar şi nu poate corespunde pentru toate ţările sau regiunile şi pentru toţi anii. Din aceste motive, calculul exact se poate efectua doar în cazul cunoaşterii vîrstei pe luni a celor decedaţi sub 1 an, ceea ce se poate obţine cu uşurinţă cînd pe buletinul de deces figurează data naşterii.

Astfel, dacă din 100 de decedaţi sub 1 an, din luna ianuarie 1963, 40 ar fi avut vîrsta de viaţă pînă la o lună, restul de 60 din cei decedaţi sub 1 an din această lună s-ar fi născut în anul 1962 . La fel şi în luna februarie 1963, dacă din 107 decese sub 1 an 53 ar fi avut vîrsta de viaţă sub 2 luni, ceilalţi 54 decedaţi s-ar fi născut în anul 1962. Efectuînd calculele în continuare, lună de lună, putem afla exact proporţia de decedaţi sub 1 an, născuţii în anul 1963 şi cei născuţi în anul 1962 şi de aici să calculăm mortalitatea infantilă, plecînd de la o proporţie justă a născuţilor vii din cei doi ani la care trebuie să raportăm decesele sub 1 an din anul 1963, în locul proporţiei amintite de 2/3 şi 1/3.

Pentru ţara noastră, această proporţie, estimată pe baza datelor privind decesele sub 1 an, nu este de 2/3 şi 1/3, ci de aproximativ 1/4 şi 3/4.

După o perioadă de 4—5 ani, acest corectiv trebuie recalculat, ţinînd seamă că după un timp mai îndelungat structura pe grupe de vîrstă a mortalităţii infantile se poate modifica, iar atunci cînd mortalitatea in-

fantilă se reduce la aproximativ 3,0%, majoritatea deceselor avînd loc în prima lună, necesitatea aplicării corectivului devine nulă.

O altă metodă de corectare a mortalității infantile este aceea de a raporta separat decesele sub 1 an din anul 1963, de exemplu, născuți în anul 1963 și separat decesele sub 1 an din 1963 ale celor născuți în 1962 conform formulei:

$$Mi_c = \left(\frac{Mi_1}{N_1} + \frac{Mi_2}{N_2}\right) \cdot 100$$

Mi_c = mortalitate infantilă corectată;
Mi_1 = decese sub 1 an, născuți în același an;
Mi_2 = decese sub 1 an, născuți în anul anterior decesului;
N_1 = născuți vii din anul respectiv;
N_2 = născuți vii din anul anterior.

Să calculăm, de exemplu, mortalitatea infantilă dintr-o regiune oarecare, unde din cei 1 200 de decedați sub 1 an, 1 000 s-ar fi născut în anul 1963 și 200 în anul 1962, iar născuții vii din acești ani ar fi fost 20 000 în 1963, respectiv 22 000 în anul 1962.

După formula clasică, mortalitatea infantilă în anul 1963, ar fi:

$$Mi = \frac{1\,200 \cdot 100}{20\,000} = 6{,}0\,\%.$$

După formula corectată prin raportarea la 2/3 născuți vii, respectiv 1/3 născuți vii, din cei doi ani, indicele ar fi diferit, și anume:

$$Mi_c = \frac{1\,200 \cdot 100}{13\,333 + 7\,333} = \frac{1\,200 \cdot 100}{20\,666} = 5{,}81\,\%.$$

Luînd ca bază de corecție 1/4 născuți vii din anul 1962 respectiv 3/4 născuți vii din anul 1963, indicele va fi:

$$Mi_c = \frac{1\,200 \cdot 100}{5\,500 + 15\,000} = \frac{1\,200 \cdot 100}{20\,500} = 5{,}85\,\%.$$

După formula corectată prin calculul separat al deceselor sub 1 an, născuți în anul 1963 și 1962, calculul va fi:

$$Mi_c = \left(\frac{1\,000}{20\,000} + \frac{200}{22\,000}\right) \cdot 100 = 5{,}91\,\%.$$

În cazul comparării mortalității infantile dintre două regiuni cu natalitate diferită sau chiar în cazul urmăririi dinamice a mortalității infantile în aceeași regiune există, din motivele anterior expuse, o eroare mai mult sau mai puțin importantă.

Este de la sine înțeles că mortalitatea sub 1 an este mult mai mare la familiile care au un număr foarte mare de copii, decît la familiile cu unul sau doi copii, ținînd seama, bineînțeles, de factorii social-economici, culturali, medico-sanitari etc.

Deci, compararea mortalității infantile dintre două regiuni cu fecunditate conjugală diferită nu este corectă. Din aceste motive este necesară corectarea mortalității infantile și din acest punct de vedere.

Aceasta se poate obține doar în cazul în care cercetarea cuprinde decesele sub 1 an, după ordinea de naștere, adică al cîtelea copil născut viu al unei mame este cel decedat sub 1 an.

Desigur că ar interesa și al cîtelea născut este cel decedat, luînd în considerare și pe cei născuți morți ; totuși, deoarece născuții morți în unele țări nu sînt bine înregistrați, se iau în considerare doar copiii născuți vii.

Mergînd la această corectare vom putea constata că mortalitatea infantilă este, în general, mai crescută la mamele cu mai mulți copii, născuți în ordinea a 5-a (al cincilea născut viu) decît la cei de ordinea întîia sau a 2-a (primul sau al doilea născut viu dintr-o familie).

Această corectare este posibilă și necesară în țările sau regiunile cu fecunditate și mortalitate infantilă ridicate.

G. DURATA MEDIE DE VIAȚĂ

Un indice care exprimă sintetic starea de sănătate a unei populații este durata medie de viață.

Probabilitatea de supraviețuire pentru fiecare an de vîrstă și durata medie de viață a unei populații se calculează pe baza tabelelor de mortalitate. Aceste tabele se calculează cu ocazia recensămintelor populației, dar se pot calcula și între recensăminte, în cazul în care se cunoaște an de an mișcarea naturală și mecanică a populației pe sexe și grupe de vîrstă.

Calculul duratei medii de viață trebuie efectuat pe țară, regiuni, pe medii și sexe. În mod asemănător se calculează durata medie de viață la diferite categorii de boli : tuberculoză, cancer, diabet sau pentru diferite profesiuni etc.

Pentru calculul duratei medii de viață este nevoie de cunoașterea populației pe ani sau grupe de vîrstă, din 5 în 5 ani, precum și mortalitatea pe ani sau pe aceleași grupe, din 5 în 5 ani. Ținînd seama de fluctuațiile mortalității pe grupe de vîrstă, de la un an la altul, se face media pe 3 ani : mortalitatea pe grupe de vîrstă din anul recensămîntului, din anul anterior și următor recensămîntului.

Luînd o populație inițială de 100 000 sau 10 000 locuitori și cunoscînd mortalitatea pe fiecare an sau grupă de vîrstă, vom putea calcula, prin scăderea celor decedați, numărul supraviețuitorilor la începutul fiecărui an de vîrstă sau la fiecare grupă de vîrstă. Acest număr de locuitori care supraviețuiesc pînă la vîrsta respectivă de 1 an sau 5 ani etc. se notează cu l_x, iar cu l_{x+1} numărul supraviețuitorilor care au împlinit $x + 1$ ani, diferența între ele corespunzînd deceselor survenite în această perioadă.

Probabilitatea de moarte corespunde întocmai acestor decese de o anumită vîrstă, raportată la populația de aceeași grupă de vîrstă și se notează cu q_x, căruia îi va corespunde deci următoarea formulă :

$$q_x = \frac{l_x - l_{x+1}}{l_x}.$$

Notînd $l_x - l_{x+1}$ cu d_x, adică decesele din anul sau grupa de vîrstă respectivă, formula devine :

$$q_x = \frac{d_x}{l_x}.$$

Durata medie a vieții corespunde probabilității medii de supraviețuire a persoanelor de o anumită vîrstă x și se notează cu e_x, iar calcularea acestui indice se face prin media aritmetică ponderată a supraviețuitorilor din an în an sau pe grupa de vîrstă pe baza formulei următoare :

$$e_x = \frac{1}{2} + \frac{l_{x+1} + l_{x+2} + \ldots + l_{x+n}}{l_x}.$$

Dăm, în continuare, tabelul XXV, cu supraviețuitorii din R.P.R. pe fiecare an de viață, luînd o populație de 100 000 locuitori.

În coloana 1 sînt înșiruiți anii (x), succesiv, de la naștere pînă la 100 de ani, peste care, în general, nu se calculează decesele pe ani de vîrstă. În coloana a 2-a se trec supraviețuitorii (l_x) la începutul anilor ce corespund coloanei 1, diferența de la un an la altul corespunzînd numărului de decese ($l_x - l_{x+1} = dx$).

În coloana a 3-a se trece indicele de mortalitate la 1 000 de locuitori, de vîrstă respectivă, iar în coloana a 4-a, cifra absolută de decese de vîrstă respectivă. Coloana a 5-a (e_x) corespunde duratei medii de viață, an de an.

Se începe cu 100 000 de locuitori, număr înscris în primul rînd al supraviețuitorilor corespunzînd vîrstei 0 (sub un an, din coloana 1, rîndul întîi). Numărul de 100 000 s-a luat în vederea ușurării calculului.

Se scad anual decesele calculate cu ajutorul indicelui de mortalitate pe fiecare an de vîrstă. Dacă se face calculul pe grupe de vîrstă din 5 în 5 ani, se ia media ponderată a anilor de vîrstă care intră în grupa respectivă. Se continuă calculele pînă la 100 de ani, deoarece statisticile nu dau, în general, date pe grupe de vîrstă ale populației decît pînă la 99 de ani sau grupa de vîrstă 95—99 de ani, cei de 100 de ani și peste fiind dați global.

Pe baza acestui tabel se trece la un calcul simplu, deși nu suficient de precis. Pentru necesitățile curente, în vederea stabilirii duratei medii de viață, se poate utiliza însă această formulă care este ușor de efectuat :

$$\left(e_x = \frac{1}{2} + \frac{l_{x+1} + l_{x+2} + \ldots + l_{x+n}}{l_x} \right)$$

Pentru o mai mare exactitate ar fi trebuit să calculăm separat durata medie de viață sub un an, pe luni. De asemenea, pentru grupele de vîrstă bătrîne ar fi trebuit efectuate și ajustări. Metoda schematică de mai sus a ținut seama doar de mortalitatea pe grupe de vîrstă, pe ani calendaristici și nu pe anii de naștere ai celor decedați. De asemenea, s-a luat în considerare numai mortalitatea pe grupe de vîrstă din anul 1956 și nu media pe 3 ani, lipsă ce este compensată, de altfel, prin faptul că, în general, acest an are indici intermediari față de anul 1955 cu indici mai mici și anul 1957 cu indici mai mari de mortalitate.

Tabelul XXV

Estimarea duratei de viaţă în R.P.R. (1956)

(calcule aproximative)

Anii	l_x	q_x ‰	q_x	e_x	Anii	l_x	q_x ‰	q_x	e_x
0	100 000	89,32	8 932	63,32	52	79 103	8,49	672	23,04
1	91 068	11,10	1 011	68,38	53	78 431	9,36	734	22,22
2	90 057	4,91	442	68,14	54	77 697	10,75	835	21,42
3	89 615	2,69	241	67,46	55	76 862	9,74	749	20,64
4	89 374	2,04	182	66,64	56	76 113	12,36	941	19,82
5	89 192	1,42	127	65,77	57	75 172	15,55	1 169	19,05
6	89 065	1,11	99	64,87	58	74 003	14,05	1 040	18,33
7	88 966	1,15	102	63,93	59	72 963	17,76	1 296	17,57
8	88 864	0,97	86	63,01	60	71 667	14,97	1 073	16,86
9	88 778	0,88	78	62,07	61	70 594	22,07	1 558	16,10
10	88 700	0,86	76	61,12	62	69 036	21,52	1 486	15,42
11	88 624	1,16	103	60,17	63	67 550	24,95	1 685	14,73
12	88 521	0,89	79	59,24	64	65 865	27,34	1 801	14,07
13	88 442	0,76	67	58,29	65	64 064	26,10	1 672	13,42
14	88 375	0,98	87	57,34	66	62 392	31,81	1 985	12,74
15	88 288	0,91	80	56,39	67	60 407	34,83	2 104	12,11
16	88 208	1,05	93	55,44	68	58 303	43,50	2 536	11,50
17	88 115	1,19	105	54,50	69	55 767	51,81	2 889	10,95
18	88 010	1,36	120	53,56	70	52 878	43,28	2 288	10,47
19	87 890	1,73	152	52,63	71	50 590	66,04	3 341	9,87
20	87 738	1,40	123	51,72	72	47 249	62,56	2 956	9,46
21	87 615	1,89	166	50,79	73	44 293	67,41	2 986	8,90
22	87 449	1,74	152	49,88	74	41 307	77,36	3 196	8,54
23	87 297	1,66	145	48,97	75	38 111	71,85	2 738	8,13
24	87 152	1,96	171	48,05	76	35 373	92,42	3 269	7,64
25	86 981	1,78	155	47,14	77	32 104	106,57	3 421	7,26
26	86 826	1,67	145	46,22	78	28 683	101,98	2 925	6,95
27	86 681	1,72	149	45,30	79	25 758	136,79	3 523	6,57
28	86 532	1,80	156	44,37	80	22 235	112,64	2 504	6,38
29	86 376	1,91	165	43,45	81	19 731	118,58	2 340	5,99
30	86 211	1,65	142	42,53	82	17 391	160,12	2 785	5,60
31	86 069	2,24	193	41,60	83	14 606	172,92	2 526	5,38
32	85 876	2,03	174	40,69	84	12 080	170,07	2 054	5,19
33	85 702	2,26	194	39,76	85	10 026	156,74	1 571	4,95
34	85 508	2,44	209	38,85	86	8 455	203,52	1 721	4,59
35	85 299	2,42	206	37,94	87	6 734	221,28	1 490	4,38
36	85 093	2,76	235	37,03	88	5 244	226,71	1 189	4,20
37	84 858	3,64	309	36,13	89	4 055	262,60	1 065	3,98
38	84 549	2,74	232	35,26	90	2 990	206,51	617	3,87
39	84 317	2,53	213	34,35	91	2 373	364,08	864	3,49
40	84 104	2,84	239	33,43	92	1 509	306,16	462	3,63
41	83 865	3,25	272	32,52	93	1 047	314,24	329	3,57
42	83 593	3,46	289	31,85	94	718	332,75	239	3,52
43	83 304	3,91	326	30,73	95	479	246,68	118	3,52
44	82 978	4,48	372	29,84	96	361	268,91	97	3,18
45	82 606	4,32	357	28,97	97	264	294,74	78	2,80
46	82 249	5,11	420	28,09	98	186	311,28	58	2,34
47	81 829	5,59	457	27,23	99	128	815,38	104	1,73
48	81 372	6,11	497	26,37	100	24	788,50	19	1,71
49	80 875	6,76	547	25,52	101	5			1,50
50	80 328	6,92	556	24,69					
51	79 772	8,39	669	23,85					

II. METODE DE STABILIRE A TENDINȚEI FENOMENELOR DEMOGRAFICE

Stabilirea tendinței în dinamica fenomenelor demografice a preocupat de mult timp nu numai statisticienii, ci și oamenii de stat, mai cu seamă în ultimele secole.

Deja John Grownt (1602—1674), primar al orașului Londra, constata o serie de aspecte cu privire la tendința natalității, mortalității și excedentului natural în orașul Londra.

La sfîrșitul secolului al XVIII-lea și în secolul al XIX-lea apar o serie de teorii reacționare cu privire la tendința în evoluția populației, printre care amintim cea a lui Malthus, care reapare și astăzi sub diverse aspecte.

Tendința de creștere în progresie geometrică a populației, conform teoriei lui Malthus, a fost combătută de Marx și Engels.

Statisticienii burghezi din ultimele decenii au răspîndit numeroase teorii, după care dezvoltarea numărului de locuitori n-ar depinde de condițiile social-economice, ci de legi fixe, biologice, ce pot fi exprimate matematic.

Amintim astfel pe Pearl, Reed, Fairchild, Bowley și alții, care, comparînd evoluția societății umane cu cea a muștelor (*Drosophila melanogaster*), au dedus o serie de formule matematice de estimare a populației viitoare.

Bazîndu-se pe asemănarea curbei dinamicii populației din ultimele decenii cu o curbă logistică, au estimat populația viitoare pe timp de un secol sau chiar mai mult.

În țara noastră, Rîmneamțu, aplicînd metodele lui Pearl și Reed, a estimat, pe baza curbei logistice, populația viitoare a Romîniei, pînă în anul 2200, presupunînd, ca și cei anterior amintiți, că evoluția populației, ca și cea a muștelor *Drosophila*, se poate încadra în formule matematice :

Suprapunerea populației observate cu cea calculată, care după Rîmncamțu „ne demonstrează și încrederea mare pe care o acordăm metodei” [1], ne demonstrează tocmai contrariul, că metoda nu este de loc aplicabilă și deci concluzia că în anul 2200 populația Romîniei devine staționară nu poate fi convingătoare.

Datele de după anul 1940 ne arată un decalaj apreciabil, față de populația calculată pentru acești ani.

De altfel, astăzi, cei mai mulți demografi, chiar și unii neomalthusieni, recunosc inutilitatea acestor metode în calcularea tendinței viitoare a populației.

Astfel, la Conferința mondială de demografie, ținută la Roma, în anul 1954, a reieșit imposibilitatea ca prin metodele existente să se poată estima tendința viitoare a populației țărilor lumii : „Perspectiva îmbunătățirilor imediate în aprecierile asupra populațiilor viitoare pentru multe populații mari, unde fertilitatea este acum ridicată, poate părea mai curînd întunecată [2]...”

---

[1] *Rîmneamțu* P. — Elemente de biometrie medicală și statistică vitală ; Ed. Institutului central de statistică, București, 1939, p. 278, al. 1.

[2] x x x — Conferința mondială de demografie, „Perspectivele creșterii și distribuției populației mondiale”, Raport preliminar, pregătit de Secretariat, Distr. General E/CN. 9/114.

... „Creşterea viitoare a populaţiei — a spus Dl. Torwquist — nu poate fi determinată complet prin tendinţele trecute şi prezente. Este posibilă o mare varietate de tipuri de evoluţie, pe baza unui complex de condiţii date..." [1]

... „Dl. L. T.Badenhorst a afirmat că el nu este de părere că metodele actualmente folosite pentru proiectările de populaţie ar putea fi considerabil îmbunătăţite într-un viitor apropiat, din cauza serioaselor lacune în cunoştinţele existente despre tipurile de comportare socială a fiinţelor omeneşti" [2].

Concluzia nu trebuie să fie negativă în ceea ce priveşte estimările tendinţei viitoare a populaţiei; ea nu trebuie să se bazeze pe utilizarea concepţiei şi metodologiei neomalthusianiştilor, pe compararea societăţii umane complexe cu cea a lumii insectelor, ci să ţină seama de structura societăţii respective, condiţiile social-economice, culturale, sanitare şi mulţi alţi factori, printre care nu trebuie să neglijăm, desigur, nici cei psihici sau biologici etc.

Pe baza analizei calitative se vor putea, de cele mai multe ori, face estimări în scopuri practice pe 1—2 decenii, cu un grad suficient de exactitate.

Ajustarea curbelor. Tendinţa fenomenelor demografice nu poate fi stabilită uneori din cauza oscilaţiilor posibile de la un an la altul. Din aceste motive, unele oscilaţii pot fi considerate în mod greşit drept tendinţe pozitive sau negative, pentru ca după cîţiva ani să ne dăm seama de aprecierea greşită.

Există o serie de metode de ajustare a curbelor, dar aplicarea lor nu trebuie să fie făcută mecanic, ci pe baza studiului esenţei fenomenului. Alteori, ajustarea curbelor nu este permisă, şi anume în cazul în care ajustarea alterează tendinţa reală a fenomenului.

Deci, ajustarea se utilizează doar în cazul în care aceasta clarifică mai bine tendinţa deja cunoscută, dar nu îndeajuns de conturată.

În cazul ajustării historiogramelor fenomenelor demografice se utilizează stabilirea valorilor medii glisante, medii formate din două sau mai multe valori alăturate. Aceasta se efectuează prin stabilirea mediilor aritmetice a trei sau maximum cinci valori alăturate.

Se utilizează, în general, media glisantă a trei valori, pentru a nu şterge unele oscilaţii importante de cunoscut, printr-o valoare medie globală, care poate ascunde acest aspect.

Dăm, în continuare, ajustarea historiogramelor natalităţii în R.P.R. între anii 1932 şi 1962, prin stabilirea valorilor medii glisante din 3 în 3 ani (graficul din fig. 10).

Se pot folosi şi alte metode de ajustare în anumite probleme, ca, de exemplu, ajustarea binomială, conform triunghiului Pascal.

---

[1] x x x — Conferinţa mondială de demografie, „Perspectivele creşterii şi distribuţiei populaţiei mondiale", Raport preliminar pregătit de secretariat, Distr. General E/CN. 9/114.

[2] x x x — Conferinţa mondială de demografie, „Perspectivele creşterii şi distribuţiei populaţiei mondiale", Raport preliminar pregătit de secretariat, Distr. General E/CN. 9/114.

... „Creşterea viitoare a populaţiei — a spus Dl. Torwquist — nu poate fi determinată complet prin tendinţele trecute şi prezente. Este posibilă o mare varietate de tipuri de evoluţie, pe baza unui complex de condiţii date..." [1]

... „Dl. L. T.Badenhorst a afirmat că el nu este de părere că metodele actualmente folosite pentru proiectările de populaţie ar putea fi considerabil îmbunătăţite într-un viitor apropiat, din cauza serioaselor lacune în cunoştinţele existente despre tipurile de comportare socială a fiinţelor omeneşti" [2].

Concluzia nu trebuie să fie negativă în ceea ce priveşte estimările tendinţei viitoare a populaţiei; ea nu trebuie să se bazeze pe utilizarea concepţiei şi metodologiei neomalthusianiştilor, pe compararea societăţii umane complexe cu cea a lumii insectelor, ci să ţină seama de structura societăţii respective, condiţiile social-economice, culturale, sanitare şi mulţi alţi factori, printre care nu trebuie să neglijăm, desigur, nici cei psihici sau biologici etc.

Pe baza analizei calitative se vor putea, de cele mai multe ori, face estimări în scopuri practice pe 1—2 decenii, cu un grad suficient de exactitate.

Ajustarea curbelor. Tendinţa fenomenelor demografice nu poate fi stabilită uneori din cauza oscilaţiilor posibile de la un an la altul. Din aceste motive, unele oscilaţii pot fi considerate în mod greşit drept tendinţe pozitive sau negative, pentru ca după cîţiva ani să ne dăm seama de aprecierea greşită.

Există o serie de metode de ajustare a curbelor, dar aplicarea lor nu trebuie să fie făcută mecanic, ci pe baza studiului esenţei fenomenului. Alteori, ajustarea curbelor nu este permisă, şi anume în cazul în care ajustarea alterează tendinţa reală a fenomenului.

Deci, ajustarea se utilizează doar în cazul în care aceasta clarifică mai bine tendinţa deja cunoscută, dar nu îndeajuns de conturată.

În cazul ajustării historiogramelor fenomenelor demografice se utilizează stabilirea valorilor medii glisante, medii formate din două sau mai multe valori alăturate. Aceasta se efectuează prin stabilirea mediilor aritmetice a trei sau maximum cinci valori alăturate.

Se utilizează, în general, media glisantă a trei valori, pentru a nu şterge unele oscilaţii importante de cunoscut, printr-o valoare medie globală, care poate ascunde acest aspect.

Dăm, în continuare, ajustarea historiogramelor natalităţii în R.P.R. între anii 1932 şi 1962, prin stabilirea valorilor medii glisante din 3 în 3 ani (graficul din fig. 10).

Se pot folosi şi alte metode de ajustare în anumite probleme, ca, de exemplu, ajustarea binomială, conform triunghiului Pascal.

---

[1] x x x — Conferinţa mondială de demografie, „Perspectivele creşterii şi distribuţiei populaţiei mondiale", Raport preliminar pregătit de secretariat, Distr. General E/CN. 9/114.

[2] x x x — Conferinţa mondială de demografie, „Perspectivele creşterii şi distribuţiei populaţiei mondiale", Raport preliminar pregătit de secretariat, Distr. General E/CN. 9/114.

În cazul ajustării binomiale, formată din trei valori alăturate, se va proceda conform formulei :

$$X = \frac{a + 2b + c}{4},$$

Fig. 10. — Metoda ajustării prin stabilirea valorilor medii glisante.

În cazul mediei binomiale, formată din patru sau cinci valori alăturate, formulele vor fi :

$$X = \frac{a + 3b + 3c + d}{8}, \text{ respectiv : } X = \frac{a + 4b + 6c + 4d + e}{16}.$$

Această metodă se utilizează atunci cînd o serie are valori centrale mai mari decît cele periferice, deci mai ales în distribuţii de frecvenţă, ca, de exemplu, dezvoltarea fizică a populaţiei.

Se mai poate folosi ajustarea prin metoda celor mai mici pătrate (parabola sau alte funcţii).

În domeniul fenomenelor demografice considerăm însă ca metodă cea mai simplă de ajustare (care altorează mai puțin, în general, tendința fenomenelor) metoda ajustării prin mediile glisante.

3. METODE DE SELECȚIE ÎN STUDIUL MORBIDITĂȚII

Dintre aspectele stării de sănătate, morbiditatea este cel mai variabil fenomen, nu numai de la un an la altul, ci și de la o lună la alta. Tocmai din aceste motive morbiditatea este un indicator sensibil al stării de sănătate. Aceasta atrage, însă, după sine necesitatea de a studia cu atenție metodologia cunoașterii morbidității populației.

Înregistrările generale ale morbidității au o serie de aspecte negative, dat fiind faptul că există lipsuri în depistarea, diagnosticarea și înregistrarea cauzelor de boală, mai ales în mediul rural. Din aceste motive, studiul general al morbidității trebuie completat cu studii parțiale, pe eșantioane reprezentative.

În cercetarea parțială în domeniul morbidității se utilizează, atît ancheta, cît și monografia. Ne oprim la prima formă de cercetare, deoarece aceasta ne permite să generalizăm rezultatele observațiilor la teritorii mari : raion, regiune, țară, în timp ce monografia poate cuprinde o comună sau o circumscripție, rezultatele ei putînd fi tipice pentru un teritoriu limitat, neputîndu-se însă generaliza pe țară sau regiune.

Metoda de selecție ne permite să estimăm, cu un oarecare procent de eroare, morbiditatea pe teritorii mari.

A. NUMĂRUL EȘANTIOANELOR

Pentru studiul morbidității pe o regiune sau țară întreagă trebuie să începem cu determinarea zonelor cu aspecte diferite de morbiditate.

Cunoscînd, pe baza datelor generale, specificul teritorial al morbidității, vom putea stabili zone cu morbiditate mare sau mică de tuberculoză, boli cardiovasculare, boli transmisibile (febră tifoidă, difterie, dizenterie etc.), boli acute ale aparatului respirator sau digestiv, reumatism etc.

Eșantioanele-martor vor fi luate, ca și în cercetările parțiale demografice, din teritoriile unde intensitatea fenomenului este mică sau medie, ținînd seama și de structura populației pe grupe de vîrstă și în primul rînd de condițiile social-economice, culturale, sanitare etc.

Stabilind astfel, spre exemplu, 10 zone diferite cu aspecte specifice de morbiditate pe țară, vom putea să ne mărginim la un număr de 180 de eșantioane total, respectiv 180 de circumscripții, ceea ce corespunde, în medie, la 10 circumscripții pe o regiune.

Desigur că acest număr de eșantioane (10 pe regiune) nu înseamnă că pentru studiul stării de sănătate pe un raion este suficient să luăm o comună, ci — ținînd seama de problemele specifice raionului respectiv — 2 circumscripții.

Reducerea relativă a numărului de eșantioane, cu cît cercetarea cuprinde teritorii mai mari, se datorește faptului că în acest caz nu cercetăm fiecare raion în parte, ci doar teritorii cu specific diferit din punctul de vedere al stării de sănătate. Astfel, de exemplu, pentru studiul gușei

într-o regiune nu este necesar să luăm în studiu toate raioanele cu gușă, ci ne mulțumim cu 2—3 circumscripții cu acest specific de morbiditate. Pentru studiul complex al stării de sănătate pe întreaga țară vom putea, deci, considera eșantionul de 180 de circumscripții ca fiind reprezentativ.

B. MĂRIMEA EȘANTIOANELOR

Numărul de locuitori din circumscripțiile sanitare amintite va fi de aproximativ 900 000. Acest eșantion este suficient de reprezentativ, deși el formează doar aproximativ 5,0% din totalul populației pe țară, fiind alese doar anumite teritorii, cu specific diferit de stare de sănătate.

Studiul pe baza eșantionului reprezentativ ne duce la cunoașterea structurii stării de morbiditate, adică a tuturor îmbolnăvirilor din sînul populației anchetate, în perioada anchetei respective.

Din aceste motive ancheta trebuie repetată din 5 în 5 ani pe aceste 180 de circumscripții medicale, pentru a putea cunoaște dinamic starea de morbiditate a populației, ceea ce, completată fiind cu dinamica morbidității generale, obținută anual din aceste circumscripții sanitare, ne permite să tragem concluzii juste pe întreaga țară asupra structurii și dinamicii morbidității.

Bineînțeles că o anchetă de o extindere atît de mare nu se face numai pe baza aspectului morbidității, ci pe baza tuturor aspectelor stării de sănătate : demografie, morbiditate, dezvoltare fizică a populației ; de asemenea și pentru aspectele privind eficiența și rentabilitatea unităților și cadrelor medico-sanitare etc.

Care sînt problemele ce pot fi cunoscute prin această cercetare ?

— În primul rînd, frecvența îmbolnăvirilor, adică intensitatea morbidității.

— În al doilea rînd, starea de morbiditate, adică totalul îmbolnăvirilor vechi și noi.

— În al treilea rînd morbiditatea specifică prin anumite boli mai frecvente sau mai puțin frecvente din anumite teritorii : tuberculoză, gușă, tifos exantematic etc. Putem cunoaște, astfel, structura pe cauze, sex, grupe de vîrstă a morbidității.

— În al patrulea rînd, se poate cunoaște mult mai aprofundat incapacitatea temporară de muncă, pe care o dau unele boli, putînd face o corelație între condițiile de muncă și sănătatea muncitorilor.

— În al cincilea rînd se poate cunoaște morbiditatea pe contingente, adică grupe de indivizi, care se prezintă o dată, de două ori sau de mai multe ori cu diferite boli acute sau cronice, îmbolnăviri noi sau vechi. De asemenea, se pot cunoaște și indivizii sănătoși, care n-au suferit de nici o boală în anul respectiv.

— În sfîrșit, în al șaselea rînd, prin sistemul de înregistrare pe buletinele anuale de morbiditate și prin repetarea din 5 în 5 ani a anchetei morbidității, se pot urmări persoanele cu anumite boli, pentru a vedea ce complicații și de ce alte boli vor suferi în decursul anilor. Acesta este un nou aspect de studiu al morbidității, și anume : morbiditatea succesivă.

Concluziile ce se pot trage în urma anchetei sînt mult mai complexe. Astfel, se va putea constata măsura în care condițiile economice-sociale și cultural-sanitare influențează morbiditatea populației. De asemenea,

se vor putea trage concluzii deosebit de prețioase cu privire la eficiența activității rețelei sanitare în lupta pentru prevenirea și combaterea morbidității.

Ținînd seama de faptul că prin cele 180 de circumscripții s-a ales un eșantion reprezentativ, să vedem în ce măsură concluziile rezultate din această anchetă se pot generaliza.

Rezultatele obținute pentru anumite boli, specifice numai anumitor teritorii (pelagră, gușă, malarie etc.), nu se pot generaliza pe țară, ci doar pe teritoriile unde există astfel de boli.

În cazul bolilor cu intensitate diferită, cu variabilitate mare pe diferite teritorii : tuberculoză, sifilis, boli acute respiratorii sau digestive, generalizarea poate fi făcută pe subgrupe de teritorii ale țării, cu același specific.

În cazul bolilor cu o mică variabilitate pe diferite teritorii ale țării, concluziile obținute se pot însă generaliza pe țară. În această categorie intră, de fapt, o bună parte dintre boli. Prima problemă ce ne interesează în prelucrarea materialului rezultat din ancheta făcută este frecvența morbidității generale, a stării de morbiditate, urmînd să cercetăm morbiditatea pe sexe, grupe de vîrstă și cauze, ocupație etc.

Problema care se pune cu această ocazie este valabilitatea numărului și a mărimii eșantionului. Aceasta se poate constata, ca și în cazul cercetărilor demografice, prin calculul valorilor medii, al dispersiunii de la medie, al erorii medii, al certitudinii diferenței valorilor la diferite eșantioane, cît și prin calcularea mărimii eșantionului necesar.

C. METODA DE APRECIERE A OMOGENITĂȚII EȘANTIONULUI ȘI A GRADULUI EI DE VARIABILITATE ÎN STUDIUL MORBIDITĂȚII

Ținînd seama de variabilitatea uneori accentuată a morbidității prin diferitele boli, este necesar, în cadrul cercetărilor de selecție, să cunoaștem, pe lîngă valorile medii, și dispersiunea față de medie, precum și gradul de eroare a valorilor medii obținute.

Să luăm, de exemplu, o anchetă efectuată pe un număr de 1 000 locuitori, în care numărul de bolnavi găsiți a fost de 150.

Media bolnavilor la o sută de anchetați este, deci, de 15,0%, acest indice exprimînd morbiditatea globală pe contingente.

Pentru a calcula gradul de omogenitate al eșantionului obținut prin selecție, aplicăm formula deviației-tip utilizată în cazul frecvențelor relative :

$$\sigma = \pm \sqrt{\frac{P \cdot (100 - P)}{N}} = \pm \sqrt{\frac{15 \cdot 85}{1\,000}} = \pm \sqrt{1{,}275} = \pm 1{,}13\%.$$

Aceasta înseamnă că față de media morbidității pe contingente, în cazul eșantionului nostru de 15,0%, dispersiunea exprimată prin 3σ este de $\pm 3{,}39$.

Calculînd coeficientul de variație, obținem această dispersiune și procentual față de media noastră de 15,0%.

$$CV = \pm \frac{\sigma \cdot 100}{Mp} = \pm \frac{1{,}13 \cdot 100}{15{,}0} = \pm \frac{113{,}0}{15{,}0} = \pm 7{,}53\%.$$

Rezultă că eşantionul ales este destul de omogen.

Pentru a putea aprecia exactitatea mediei ca măsurătoare a valorii centrale, aplicăm formula erorii medii şi procentul erorii medii :

$$e\,M = \pm \frac{\sigma}{\sqrt{N}} = \pm \frac{1{,}13}{\sqrt{1000}} = \pm \frac{1{,}13}{31{,}62} = \pm\, 0{,}036.$$

$$e\,M\,\% = \pm \frac{0{,}036 \cdot 100}{15{,}0} = \pm \frac{3{,}6}{15{,}0} = \pm\, 0{,}24\,\%.$$

Eroarea medie este mică, reprezentînd 0,24% din valoarea mediei noastre de 15,0%. De aici rezultă că proporţia de 15,0% bolnavi caracterizează cu precizie realitatea în cazul eşantionului luat. Luînd chiar triplul erorii medii, obţinem $\pm$ 0,11, ceea ce indică precizia valorii mediei noastre.

D. METODA DE CALCUL A CERTITUDINII DIFERENŢEI DINTRE FRECVENŢELE MORBIDITĂŢII A DOUĂ EŞANTIOANE

În studiul morbidităţii, în cadrul cercetărilor de selecţie, trebuie, de cele mai multe ori, să comparăm indicii între ei, atît în cazul a două eşantioane diferite rezultate din aceeaşi cercetare, cît şi în cazul a două eşantioane obţinute din două anchete consecutive pentru aceeaşi populaţie.

Problema ce se pune este de a verifica în ce măsură diferenţele dintre indici sînt semnificative, certe, sau sînt doar întîmplătoare, datorită faptului că eşantioanele n-au fost alese după aceleaşi criterii.

Să considerăm, de exemplu, morbiditatea pe contingente (persoane bolnave la o sută de anchetaţi) fiind de 50,0% la o populaţie anchetată de 10 000 locuitori, iar la un alt eşantion, pe un număr identic de locuitori, o morbiditate pe contingente de 45,0%.

Prin metoda stabilirii certitudinii diferenţei putem să ne dăm seama dacă diferenţa dintre cei doi indicatori ai morbidităţii pe contingente este sau nu semnificativă.

Calculăm întîi deviaţiile-tip şi erorile medii la cele două valori medii (50,0% şi 45,0%).

$$\sigma_1 = \pm \sqrt{\frac{P_1 \cdot (100 - P_1)}{N_1}} = \pm \sqrt{\frac{50\,(100 - 50)}{10\,000}} = \pm \sqrt{\frac{2\,500}{10\,000}} = \pm\, 0{,}500.$$

$$\sigma_2 = \pm \sqrt{\frac{P_2 \cdot (100 - P_2)}{N_2}} = \pm \sqrt{\frac{45\,(100 - 45)}{10\,000}} = \pm \sqrt{\frac{2\,475}{10\,000}} = \pm\, 0{,}497$$

$$e\,M_1 = \pm \frac{\sigma_1}{\sqrt{N_1}} = \pm \frac{0{,}500}{\sqrt{10\,000}} = \pm \frac{0{,}500}{100} = \pm\, 0{,}005.$$

$$e\,M_2 = \pm \frac{\sigma_2}{\sqrt{N_2}} = \pm \frac{0{,}497}{\sqrt{10\,000}} = \pm \frac{0{,}497}{100} = \pm\, 0{,}00497.$$

Avînd datele celor două erori medii, putem aplica formula de calcul a erorii diferenței celor două medii (50,0 și 45,0%).

$$e\,D = \pm\sqrt{e\,M_1^2 + e\,M_2^2} = \sqrt{0{,}005^2 + 0{,}00497^2} = \pm\,0{,}007$$

Luînd triplul acestei erori a diferenței obținem ±0,021 mult mai scăzut decît diferența dintre cele două valori medii : 50,0—45,0% = = 5,0%, ceea ce denotă că diferența dintre aceste valori medii este semnificativă și nu întîmplătoare.

E. METODA DE CALCUL A EȘANTIONULUI REPREZENTATIV

În exemplul anterior, luînd două eșantioane de cîte 10 000 locuitori, am dedus, prin calcul, certitudinea diferenței cu o mare probabilitate. Totuși, acest calcul nu ne indică în mod exact dacă cele două eșantioane nu au fost prea mari, fiind suficient să luăm poate jumătate din populația respectivă și să obținem , prin eforturi mult mai mici, rezultate tot atît de exacte în stabilirea morbidității pe contingente. În acest scop utilizăm formula cunoscută a eșantionului reprezentativ, calculînd, mai întîi, dispersiunea diferenței, din care se deduce apoi formula eșantionului reprezentativ :

$$\sigma D = \pm\sqrt{\frac{P_1(100 - P_1) + P_2(100 - P_2)}{N}} = \pm\sqrt{\frac{50\,(100 - 50) + 45\,(100 - 45)}{10\,000}} =$$

$$= \pm\sqrt{\frac{4\,975}{10\,000}} = \pm 0{,}70.$$

Triplul deviației-tip a diferenței celor două valori medii de morbiditate pe contingente este de ±2,10 (3.0,70 = 2,10), mai puțin de jumătatea diferenței morbidităților pe contingente, care este de 5,0% (50,0%— —45,0% = 5,0%).

Eșantioanele noastre sînt mari, după cum rezultă din compararea indicatorilor (5,0 > 2,10).

Cît de mic ar putea fi eșantionul, cu condiția ca el să mai poată fi reprezentativ?

Acesta se calculează, dacă în loc de σD = ± 0,70 am fi luat un σD, care triplat ar fi sub 5,0, cît este diferența dintre cele două valori medii. Acest σD ar putea fi 1,66, care triplat dă 4,98 (deci sub 5,0).

Aplicăm formula eșantionului reprezentativ, cu utilizarea unui σD = ±1,66 în loc de ±0,75.

$$N = \frac{P_1(100 - P_1) + P_2(100 - P_2)}{\sigma D^2} = \frac{2\,500 + 2\,475}{1{,}66^2} = 1\,802.$$

Eșantionul reprezentativ ar fi fost 1 802, număr suficient de cazuri care să demonstreze, cu o probabilitate de 99,73%, certitudinea diferenței celor doi indicatori ai morbidității pe contingente.

F. METODA DE CORECTARE A MORBIDITĂŢII

Ca şi în cazul fenomenelor demografice, morbiditatea generală sau morbiditatea cu incapacitate temporară de muncă ne pot induce în eroare, în cazul în care se compară doi indici medii globali, neţinînd seamă de distribuţia lor pe sexe, grupe de vîrstă etc.

Oscilaţiile în cadrul dinamicii morbidităţii fiind cu mult mai mari decît cele privind dinamica fenomenelor demografice, se înţelege de ce este necesar ca indicii de morbiditate să fie corectaţi prin diferite metode.

În ceea ce priveşte sexul, morbiditatea diferă nu numai pe anumite cauze de îmbolnăvire, specifice unuia sau celuilalt sex, dar şi în diferite boli, cum ar fi : bolile hepatobiliare, boala hipertonică, febra tifoidă etc.

Diferenţe mari există mai cu seamă în legătură cu grupele de vîrstă, morbiditatea fiind mult mai mare la sugari şi la copiii mici, pe de o parte, şi la bătrîni, pe de alta.

Această diferenţă este mult mai accentuată în anumite boli, la copiii mici predominînd bolile acute respiratorii, de nutriţie şi digestie, boli infecto-contagioase, iar la bătrîni, diferite boli cronice : bolile cardio-vasculare, tumori, boli renale etc.

Desigur că structura morbidităţii ca şi a mortalităţii se schimbă o dată cu îmbunătăţirea stării social-economice, cultural-sanitare şi cu progresul ştiinţei medicale. Astfel este cazul mortalităţii prin tuberculoză, care avea cea mai mare frecvenţă la adolescenţi, cu cîţiva ani în urmă, pentru ca în prezent mortalitatea cea mai mare să se înregistreze la adulţi şi vîrstnici.

Ţinînd seama de aceste considerente, ori de cîte ori vom compara structura şi dinamica morbidităţii, va trebui s-o analizăm prin prisma tuturor elementelor, dintre care cele mai importante au fost amintite anterior.

Corectarea cea mai importantă, care trebuie să stea în mijlocul atenţiei, în cazul analizei morbidităţii, este corectarea pe sexe şi în special pe grupe de vîrstă.

S-ar putea ca morbiditatea, cu sau fără incapacitate temporară de muncă dintr-o întreprindere, să crească în cîţiva ani numai din motivul că s-au schimbat grupele de vîrstă, crescînd media vîrstei salariaţilor de la 26 de ani la 36 de ani.

Studiul morbidităţii pe mediul urban, rural, regiuni, sexe şi grupe de vîrstă se poate face cu precizie în special în anii recensămintelor, cunoscînd, atît populaţia, cît şi îmbolnăvirile pe sexe, grupe de vîrstă. În anii dintre recensăminte, calculul se poate face doar cu aproximaţie. Metoda de calcul al morbidităţii corectate este identică cu cea a mortalităţii corectate, obţinută prin raportarea la populaţia standard.

G. TENDINŢE ÎN DINAMICA MORBIDITĂŢII

Ţinînd seama de cele arătate la fenomenele demografice despre dificultatea de a stabili tendinţele pe o perioadă mai îndepărtată, amintim că aceasta este o problemă cu mult mai dificilă în domeniul morbidităţii.

Dacă în domeniul fenomenelor demografice (natalitate, mortalitate generală sau infantilă etc.) nu putem să ne aşteptăm la variaţii importante, în decursul a 3 sau 4 ani, în domeniul morbidităţii putem avea,

după o scădere continuă an de an, o creştere bruscă şi accentuată de la un an la altul, sub influenţa diverşilor factori ce n-au putut fi prevăzuţi. Variabilitatea este, deci, cu mult mai accentuată în cazul morbidităţii decît în cazul fenomenelor demografice.

Ar fi totuşi nejust să afirmăm că nu s-ar putea şi în cazul morbidităţii stabili tendinţele de evoluţie. Această previziune pe bază de calcule este posibilă şi bineînţeles este deosebit de necesară în vederea prevenirii şi combaterii morbidităţii şi mortalităţii.

Tendinţele în evoluţia morbidităţii şi mortalităţii se pot stabili, în primul rînd, din evoluţia în ultimele decenii a acestora în diferite ţări. Din acest punct de vedere se pot stabili chiar tipuri de ţări cu morbidităţi şi mortalităţi specifice, după cum în cadrul aceleiaşi ţări se pot stabili regiuni cu specific de morbiditate sau mortalitate.

Această morbiditate specifică în realitate nu este tipică pentru o anumită ţară sau regiune decît într-o perioadă anumită, în strînsă legătură cu evoluţia condiţiilor social-economice, culturale, igieno-sanitare etc.

Nu trebuie să neglijăm o serie de factori, cum ar fi clima, care determină o serie de tipuri de morbiditate, care pot fi eliminate şi ele, o dată cu dezvoltarea metodelor de prevenire şi combatere a acestor boli. Amintim, în acest sens, în ţara noastră, doar bolile acute ale aparatului respirator în anotimpul de iarnă şi începutul primăverii, în scădere importantă în ultimii ani.

Nu trebuie să neglijăm nici o serie de boli legate de sex şi grupe de vîrstă care pot determina un anumit specific de morbiditate sau mortalitate. Este clasică, în acest sens, creşterea morbidităţii şi mortalităţii în bolile cardiovasculare la grupele vîrstnice ale populaţiei şi de aici creşterea acestor afecţiuni în ţările cu natalitate scăzută, deci cu o proporţie crescută de populaţie vîrstnică. Şi în aceste boli, însă, măsurile de profilaxie pot preveni şi reduce mult morbiditatea şi mortalitatea.

Desigur că în bolile şi decesele datorite acestei afecţiuni nu trebuie să ne rezumăm în a incrimina drept cauză doar structura pe vîrstă a populaţiei, deoarece o serie de alţi factori, care pot duce la lezarea timpurie a aparatului cardiovascular, pot influenţa, în mare parte, mortalitatea sau morbiditatea crescută prin aceste cauze.

În cadrul morbidităţii specifice pe tipuri de ţări amintim ţările cu o morbiditate mare prin bolile infectocontagioase şi parazitare, apoi tipul ţărilor care au reuşit să domine aceste boli, reducîndu-le în mare măsură.

Aici, în aceste ţări, ajung să domine bolile acute ale aparatului respirator şi digestiv, îndeosebi la sugari şi copiii mici.

Totuşi o serie de boli transmisibile, cum ar fi gripa, reuşesc, în unii ani, să ducă la epidemii extinse, ceea ce face ca tipul de structură a morbidităţii să încline din nou spre predominarea bolilor infecto-contagioase.

Amintim, de asemenea, tipul ţărilor cu morbiditate mare în boli reumatice, cardiovasculare, boli de nutriţie etc. care domină structura pe cauze a morbidităţii în unele ţări din Occident, legate nu numai de scăderea grupelor tinere de populaţie, ci şi de o serie de alte cauze, printre care cele care duc la uzura psihosomatică sînt printre cele mai importante.

Este greşită însă judecarea schematică, simplistă a cauzelor care determină structura şi dinamica morbidităţii şi este necesar să cercetăm,

în mod complex, cauzele care determină, respectiv care influențează, tendința în evoluția ei.

Tendința de evoluție oarecum mai lentă a morbidității pe linia scăderii bolilor transmisibile, apoi a bolilor acute ale aparatului respirator, digestiv și a tuberculozei a fost și este modificată din timp în timp de modificări în salturi cu ocazia descoperirii și aplicării în masă a unui nou medicament sau măsuri de asanare a condițiilor de mediu (amintim la noi scăderea mortalității prin tuberculoză sau a morbidității prin malarie, febră tifoidă, difterie etc.).

Un rol de cea mai mare importanță îl au, în sfîrșit, accesibilitatea și adresabilitatea populației la o asistență medicală calificată și specializată, deci numărul de cadre și unități medico-sanitare, repartiția lor teritorială, pregătirea lor profesională, cît și nivelul de cultură sanitară a populației. În cadrul accesibilității, o importanță de bază o prezintă și organizarea ocrotirii sănătății, dacă aceasta se ghidează după principiile socialiste, asistența populației fiind o problemă de Stat sau dacă ea este lăsată pe seama instituțiilor de binefacere, sau a medicinii mercantile, asistența devenind o problemă de profit, ca în țările capitaliste.

De toate aceste aspecte complexe trebuie să ținem seama, atunci cînd stabilim tendința în dinamica și structura morbidității și mortalității. Vom putea, deci, stabili tendințele viitoare ale morbidității, analizînd tendințele din trecut și din prezent ale fiecărei boli în parte, pe grupe de vîrstă, pe diferite teritorii etc.

Vom putea astfel stabili, pe cîțiva ani (5—10 ani, în funcție de fiecare grup de boli), tendința viitoare de evoluție, ca, de exemplu, evoluția morbidității prin bolile cardiovasculare pe 10 ani sau evoluția tuberculozei și a hepatitei epidemice pe aproximativ 5 ani.

În cazul morbidității nu vom utiliza metoda ajustării decît cu mare grijă și doar pe cîțiva ani. Vom putea și este necesar însă să utilizăm, pentru stabilirea tendinței, metoda ajustărilor prin stabilirea valorilor medii glisante sau ajustarea binominală etc., ca și în cazul fenomenelor demografice, dar numai pe o perioadă de maximum 3 ani.

4. CERCETAREA DE SELECȚIE ÎN DOMENIUL DEZVOLTĂRII FIZICE A POPULAȚIEI

Indicatorii somatometrici sînt de mare importanță în analiza stării de sănătate, reflectînd dezvoltarea fizică armonioasă a populației.

Dezvoltarea fizică a populației reflectă, în mod sintetic, starea de sănătate a populației, condițiile social-economice, culturale, igieno-sanitare, morbiditatea populației etc. Îmbunătățirea indicatorilor de dezvoltare fizică denotă întotdeauna îmbunătățirea stării de sănătate și invers.

Cercetarea dezvoltării fizice a populației utilizează întotdeauna metoda cercetării parțiale pe baza eșantionului reprezentativ. Acesta se face alegînd copii și tineri pînă la 25 de ani de obicei, deci perioada de dezvoltare a organismului, deși din punct de vedere științific este interesant să se cunoască acești indicatori pentru toate vîrstele. Pentru studiul stării de sănătate este însă suficient să cunoaștem indicatorii somato-

metrici pînă la 20—25 de ani, care în realitate reflectă dezvoltarea fizică a întregii populații.

Există o serie de date care cuprind nu numai un eşantion reprezentativ, ci uneori întreaga grupă de vîrstă. Amintim, în acest sens, măsurătorile la recruți, la nou-născuți, copiii din creşă, şcoli etc. În acest fel se adună un material foarte vast, pe baza căruia se poate studia dinamica dezvoltării fizice a populației.

Studiul acesta trebuie să fie completat de un material cules pe bază de cercetare riguros ştiințifică. Aceasta necesită o metodologie uniformă, conducere unică a cercetării, controlul pe teren al datelor înregistrate, o anumită perioadă a măsurătorii şi, bineînțeles, o metodologie uniformă de prelucrare şi interpretare.

Din aceste motive este necesar să se recurgă din 5 în 5 ani sau chiar din 10 în 10 ani la o anchetă, alegîndu-se un eşantion suficient de mare pentru a se putea trage concluzii ştiințifice care se pot generaliza la populația din întreaga regiune sau din toată țara.

Pe baza unei documentații temeinice, în prealabil, se trece la alegerea numărului şi mărimii eşantionului.

A. NUMĂRUL EŞANTIONULUI

Deoarece în studiul dezvoltării fizice a populației sîntem interesați să cunoaştem diferitele tipuri specifice, este necesar ca numărul de eşantioane să fie suficient de numeros pentru a putea să reflecte, atît situația medie pe țară, cît şi situația specifică pe diferite teritorii.

Vom ține seama de condițiile geografice (relief, climă), condițiile istorice de dezvoltare specifică a unor teritorii, condițiile social-economice, cultural-sanitare, de diferite aspecte specifice de morbiditate, mortalitate, natalitate, migrațiuni etc.

Trebuie luate eşantioane din mediul rural şi urban, deoarece se ştie că există diferențe între aceste medii din punctul de vedere al indicilor de dezvoltare fizică ; de asemenea, trebuie luate eşantioane separate pentru ambele sexe şi bineînțeles pe grupe de vîrstă.

În ceea ce priveşte vîrsta, trebuie să menționăm că ea este deosebit de importantă, de la o vîrstă la cealaltă existînd diferențe uneori foarte mari. Din această cauză, sub vîrsta de 1 an, vom grupa eşantioanele din lună în lună, iar între 1 an şi 3 ani din 3 în 3 luni. Între 3 şi 6 ani vom lua eşantioane din jumătate în jumătate de an, iar de la 7 ani în sus putem lua din an în an. În ceea ce priveşte sexele, vom lua eşantioane separate pentru fiecare vîrstă şi sex în parte.

În consecință, numărul de eşantioane va trebui să fie foarte mare, şi anume : aproximativ 15—20 de teritorii, împărțite fiecare pe urban şi rural, 2 sexe şi 40 de grupe de vîrstă, 3—4 categorii sociale. În total vor trebui luate aproximativ 6 800 de eşantioane, în cazul unui studiu pe întreaga țară.

Specificăm că la ancheta privind dezvoltarea fizică a populației din anii 1950 şi 1957, organizată de Ministerul Sănătății şi Prevederilor Sociale, au fost alese, pe bază de studii, 11 regiuni împărțite pe 20 de subregiuni (graficul din fig. 11).

În anul 1950 au fost anchetați numai copiii între 3 şi 18 ani, în timp ce la ancheta din 1957 copiii de la 0 la 18 ani. La gruparea

materialului s-a mers la vîrsta din lună în lună sub 1 an și din 3 în 3 luni la vîrste de 1 și 2 ani. Începînd din al treilea an de viață s-a grupat materialul pe vîrste din an în an. Toate aceste grupe de vîrstă au fost clasificate separat pe sexe.

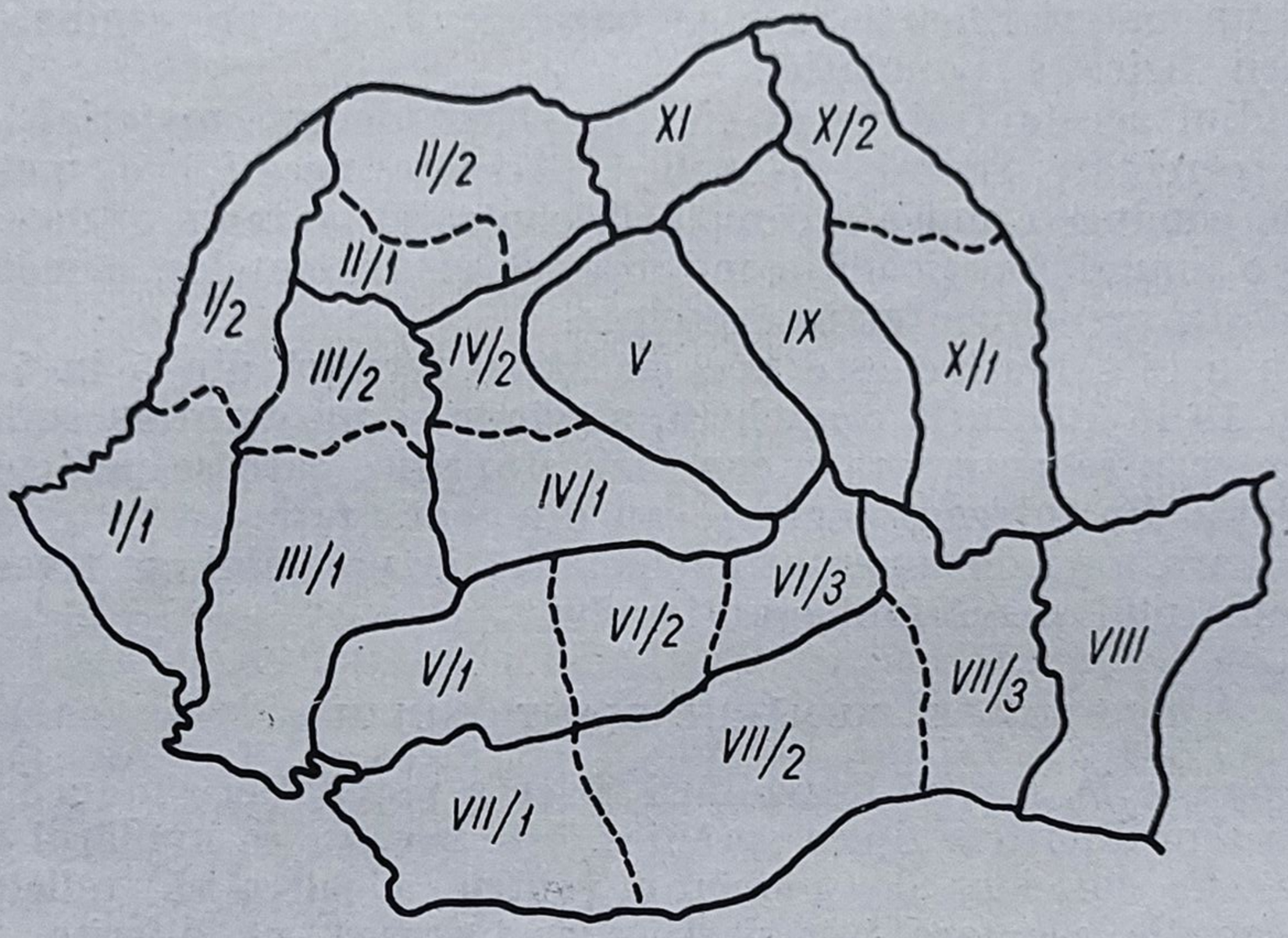

Fig. 11. — Zonele și subzonele de cercetare a dezvoltării fizice a copiilor.

Au fost luate eșantioane, atît din mediul rural, cît și din cel urban, și anume pe categorii de orașe : orașe mai mari și mai mici, mai industrializate sau mai puțin industrializate etc.

Cu privire la eșantionul martor în problema dezvoltării fizice a populației, acesta se pune într-un mod cu totul diferit. Fiecare dintre regiuni sau subregiuni se compară între ele neexistînd una care să servească ca etalon pentru celelalte.

În cazul repetării anchetei, comparația trebuie însă făcută pentru a putea vedea dinamica dezvoltării fizice a populației. Doar în acest caz am putea vorbi, în realitate, de eșantioane martor, adică cele de la ancheta anterioară cu care comparăm ancheta recentă.

În realitate, însă, nu se poate vorbi de un eșantion martor în cazul dezvoltării fizice a populației. De altfel s-a arătat, în capitolele anterioare, că nici în domeniul fenomenelor de demografie sau morbiditate nu putem vorbi de eșantioane martor, cum există și trebuie să existe în cercetările de laborator.

B. MĂRIMEA EȘANTIONULUI

Avînd în vedere numărul mare de eșantioane, aproximativ 6 800, se înțelege că și mărimea totală a eșantioanelor va fi mare. Luînd ca număr minim pentru un eșantion 40 de indivizi, considerat ca eșantion minimal, mărimea totală va ajunge la aproximativ 270 000.

Ancheta de o astfel de amploare trebuie organizată central, cu echipe bine și uniform instruite, avînd o conducere centrală și regională, instrumentar de măsurătoare de mare precizie (greutatea trebuie cîntărită pînă la mărimi de sute de grame, iar înălțimea, înălțimea șezînd, circumferința toracică și craniană etc. pînă la milimetri).

Nu ne oprim însă la metoda anchetei în teren și la pregătirea ei, ci ne rezumăm la metodele de statistică matematică, care se pun în legătură cu studiul dezvoltării fizice a populației.

Fișele în care se înregistrează datele somatometrice sînt centralizate și prelucrate uniform. Prelucrarea trece prin mai multe faze : verificarea datelor înregistrate în fișe și codificarea pentru mașinile statistice. Fiecare dintre aceste două faze este din nou controlată pentru a avea siguranța unei prelucrări exacte.

Prin codificarea fișelor s-a și decis posibilitatea grupărilor ulterioare a caracteristicilor esențiale.

Prin prelucrarea mecanică la mașinile statistice se obțin o serie de tabele cu datele somatometrice (înălțimea, înălțimea șezînd, circumferința toracică, circumferința craniană, greutatea), repartizate pe regiuni, medii, sexe, grupe de vîrstă etc.

Pentru a putea interpreta aceste date deosebit de bogate, recurgem la o serie de indicatori calculați prin metoda statisticii matematice, și anume : valorile medii (media aritmetică ponderată), indicatorii de dispersiune de la medie (deviația standard și coeficientul de variație), eroarea medie, procentul erorii medii și eroarea probabilă a mediei, indicatorii de corelație, indicatorii de regresie etc.

Se va putea, de asemenea, calcula certitudinea diferenței valorilor medii ale diferitelor eșantioane, obținute la aceeași anchetă și în special la anchetele efectuate în diferiți ani, pentru a aprecia, în mod just, dinamica dezvoltării fizice.

Vom putea, de asemenea, calcula necesarul mărimii eșantionului reprezentativ, pentru a vedea dacă eșantionul global a fost sau nu suficient sau eventual mult peste necesar.

Dăm ca exemplu înălțimea copiilor de 10 ani, de sex masculin, fii de muncitori, dintr-o anumită regiune A.

Distribuția cazurilor este cea arătată în tabelul XXVI.

C. CALCULUL VALORII MEDII A EȘANTIONULUI

În prima etapă calculăm valoarea centrală : media aritmetică ponderată. Pentru aceasta alegem o medie ipotetică, pe care o socotim cea mai apropiată de realitate, deci cea care corespunde unor frecvențe mai mari, deoarece frecvențele, dacă eșantionul este omogen și suficient de mare, sînt aranjate conform curbei ideale, frecvențele mari situîndu-se central, corespunzînd medianei, care, de cele mai multe ori, se află în preajma mediei aritmetice ponderate.

În cazul nostru am ales ca medie aritmetică, în mod ipotetic, înălțimea de 131 cm corespunzătoare unei frecvențe mari de 76 de cazuri.

Tabelul XXVI

Înălțimea copiilor din regiunea A., în vîrstă de 10 ani, sex masculin, fii de muncitori

Varianta (în cm) $= X$	Frecvența $= n$	Deviația de la medie $= X'$	Deviația de la medie, înmulțită cu frecvența $= X'.n$	Deviația medie la pătrat $= (X')^2$	Deviația medie la pătrat, înmulțită cu frecvența $= (X')^2.n$
1	2	3	4	5	6
126	2	−5	−10	25	50
127	4	−4	−16	16	64
128	10	−3	−30	9	90
129	14	−2	−28	4	56
130	36	−1	−36	1	36
			−120		
131	76	0	0	0	0
132	32	+1	+32	1	32
133	18	+2	+36	4	72
134	8	+3	+24	9	72
135	4	+4	+16	16	64
136	4	+5	+20	25	100
137	2	+6	+12	36	72
			+140		
	$N = 210$;		$\Sigma X'.n = 20$;		$\Sigma(X')^2.n = 708$

Se trece la calculul mediei aritmetice ponderate prin metoda momentelor (momentul 1):

$$M_1 = \frac{\Sigma X' \cdot n}{N} = \frac{\text{coloana a 4-a}}{\text{coloana a 2-a}} = \frac{20}{210} = 0{,}095.$$

Media aritmetică ponderată va fi egală cu media ipotetică plus momentul 1:

$$M_p = 131 + 0{,}095 = 131{,}1 \text{ cm.}$$

Se poate vedea că media ipotetică a fost bine aleasă, ea fiind aproape identică cu media aritmetică ponderată, aflată prin calcul.

Deci media înălțimii copiilor de 10 ani, sex masculin, fii de muncitori din regiunea A, este de 131,1 cm.

D. CALCULUL DISPERSIUNII DE LA MEDIE

Pentru calculul dispersiunii, prin indicele deviației-tip, calculăm, în prealabil, momentul al II-lea:

$$M_2 = \frac{\Sigma (X')^2 \cdot n}{N} = \frac{\text{coloana a 6-a}}{\text{coloana a 2-a}} = \frac{708}{210} = 3{,}37.$$

Cu ajutorul momentului 1 și al II-lea calculăm apoi deviația-tip:

$$\sigma = \pm \sqrt{M_2 - (M_1)^2} = \pm \sqrt{3{,}37 - 0{,}095^2} = \pm \sqrt{3{,}361} = \pm 1{,}84$$

Se poate constata că eşantionul este omogen, deoarece în intervalul dintre media ponderată a înălţimii şi $\pm 3\ \sigma$ se cuprinde aproape totalitatea cazurilor din eşantionul nostru de 210 copii :

$$M_p \pm 3\sigma = 125{,}58 \text{ respectiv } 136{,}62 \text{ cm.}$$

Pentru a putea estima procentul dispersării se calculează coeficientul de variaţie :

$$cV\cdot = \frac{\sigma\cdot 100}{M_p} = \frac{1{,}84\cdot 100}{131{,}1} = 1{,}40\,\%.$$

Putem afirma, avînd în vedere acest coeficient de variaţie foarte mic, că eşantionul este omogen, bine ales, deci reprezentativ pentru regiunea întreagă.

E. CALCULUL ERORII MEDII

Ca să vedem în ce măsură media ponderată calculată pe eşantionul nostru corespunde sau se îndepărtează de la media absolută, ideală, care ar fi fost obţinută în cazul în care am fi măsurat toţi copiii din această categorie din regiune, calculăm eroarea medie a mediei şi eroarea probabilă :

$$eM = \pm\frac{\sigma}{\sqrt{N}} = \pm\frac{1{,}84}{\sqrt{210}} = \pm\frac{1{,}84}{14{,}49} = \pm\ 0{,}13 \text{ cm.}$$

Cu o probabilitate de 99,73%, putem afirma că media absolută se află între limitele mediei ponderate $\pm\ 3\ eM$, adică între 130,71 şi 131,49 cm, ceea ce este o limită destul de strînsă :

$$Mp \pm 3\ eM = 131{,}1 \pm 3\cdot 0{,}13 = 130{,}71 \text{ respectiv } 131{,}49.$$

Pentru a putea estima procentual valoarea erorii medii a mediei facem următorul calcul :

$$eM\% = \pm\frac{eM\cdot 100}{Mp} = \pm\frac{0{,}13\cdot 100}{131{,}1} = \pm 0{,}099\,\%$$

Dacă luăm triplul procentului erorii medii (0,3%) constatăm că ea este deosebit de mică (sub 3,0%), deci media ponderată ce a fost calculată de noi (131,1 cm) caracterizează bine valoarea medie a înălţimii întregului eşantion.

Unii autori consideră mai exact de a utiliza, nu eroarea medie a mediei, ci eroarea probabilă, care se calculează în felul următor :

$$ep = \pm eM\ 0{,}6745 = \pm 0{,}09 \text{ cm.}$$

Probabilitatea ca media absolută să fie cuprinsă între intervalul dintre media ponderată empirică şi de patru ori eroarea probabilă (131,1 $\pm 0{,}36 = 130{,}74$ respectiv 131,46) este de 99,3%.

F. CORELAŢIA ÎN DOMENIUL DEZVOLTĂRII FIZICE A POPULAŢIEI

Este necesar ca la sfîrşitul tuturor calculelor privind indicii somatometrici să facem o serie de corelări între aceşti indici.

Un indicator important de corelaţie este raportul dintre înălţime şi greutate sau dintre înălţimea şezînd şi greutate.

Indicatorul sintetic care exprimă aceste raporturi de interdependenţă dintre două sau mai multe fenomene este coeficientul de corelaţie.

El se calculează, după cum s-a arătat şi în capitolul I, pe baza tabelului de corelaţie. Vom reda doar pe scurt calculul coeficientului de corelaţie.

Să luăm exemplul privind înălţimea copiilor de 10 ani,sex masculin, fii de muncitori, din regiunea A, corelată cu greutatea (tabelul XXVII).

Tabelul XXVII

Înălţimea şi greutatea copiilor din regiunea A, în vîrstă de 10 ani, sex masculin, fii de muncitori

		x = greutatea în kg												
		24	25	26	27	28	29	30	31	32	33	34	35	Suma $y=\Sigma y$
		−5	−4	−3	−2	−1	0	1	2	3	4	5	6	
126	−5	1_{25}	1_{20}											2
127	−4		3_{16}	1_{12}										4
128	−3			5_{9}	3_{6}	2_{3}								10
129	−2			4_{6}	6_{4}	3_{2}	1_{0}							14
130	−1				5_{2}	18_{1}	8_{0}	4_{1}	1_{2}					36
131	0				3_{0}	14_{0}	40_{0}	12_{0}	5_{0}	2_{0}				76
132	1					5_{1}	4_{0}	17_{1}	4_{2}	2_{3}				32
133	2							9_{2}	8_{4}	1_{6}				18
134	3							3_{3}	2_{6}	3_{9}				8
135	4									2_{12}	2_{16}			4
136	5										2_{20}	1_{25}	1_{30}	4
137	6											2_{30}		2
Suma $x=\Sigma x$		1	4	10	17	42	53	45	20	10	4	3	1	210

Din tabel rezultă corelaţia strînsă ce există între înălţime şi greutate, sensul acestei corelaţii fiind pozitiv, adică o dată cu creşterea înălţimii creşte şi greutatea şi invers.

Se poate constata că printre cei 4 copii de o înălţime de 127 cm există 3 copii avînd greutatea de 25 kg şi un copil cu greutatea de 26 kg.

De asemenea, printre cei trei copii cu o greutate de 34 kg, un copil are înălţimea de 136 cm, iar 2 copii, de 137 cm.

Cele două totaluri sau sume sînt egale, fiind vorba de un număr de 210 copii. În tabel, prin x s-a notat greutatea, iar prin y, înălţimea.

Se poate vedea, în mod sugestiv, că spre centrul tabelului, frecvențele care corespund înălțimii și greutății sînt cele mai mari, ceea ce înseamnă că există o corelație puternică, strînsă, între greutatea și înălțimea acestui lot de copii.

Pe baza tabelului de corelație se poate trece la calculul coeficientului de corelație. În acest scop, în tabel, s-au ales valorile medii de înălțime și greutate, adică media ipotetică, pentru înălțime : 131 cm, iar pentru greutate : 29 kg.

Pentru calculul coeficientului de corelație se face produsul dintre deviațiile de la medie și frecvențele corespunzătoare din fiecare pătrățel din tabelul de corelație, după care facem suma totală a acestor produse, ținînd seama de semnul algebric al valorilor (pătratul 1 și 4 are semnul plus, pătratul 2 și 3 semnul minus).

Primul pătrat (+)		Al doilea pătrat (−)	Al treilea pătrat (−)	Al patrulea pătrat (+)	
1 . 25 = 25	2 . 3 = 6	4 . 1 = 4	5 . 1 = 5	17 . 1 = 17	3 . 9 = 27
1 . 20 = 20	4 . 6 = 24	1 . 2 = 2		4 . 2 = 8	2 . 12 = 24
3 . 16 = 48	6 . 4 = 24	− 6	− 5	2 . 3 = 6	2 . 16 = 32
1 . 12 = 12	3 . 2 = 6			9 . 2 = 18	2 . 20 = 40
5 . 9 = 45	5 . 2 = 10			8 . 4 = 32	1 . 25 = 25
3 . 6 = 18	18 . 1 = 18			1 . 6 = 6	1 . 30 = 30
	+256			3 . 3 = 9	2 . 30 = 60
				2 . 6 = 12	
					+346

Suma celor patru pătrate va fi :

$$+ 256 - 6 - 5 + 346 = 591$$

Această valoare (591) reprezintă deviația de la media ipotetică aleasă de noi a celor două însușiri : înălțimea și greutatea, înmulțite cu frecvențele respective.

Pentru a afla deviațiile de la mediile ponderate necesare calculului coeficientului de corelație, precum și deviațiile-tip, utilizăm metoda momentelor care se calculează pentru fiecare însușire în parte (înălțime și greutate) pe baza tabelelor de distribuție.

Calculele pentru înălțime le cunoaștem din paginile anterioare, și anume :

$M_1 =$ 0,1 cm (de fapt 0,095 cifră rotunjită la 0,1).
$M_v =$ 131,1 cm.
$\sigma = \pm$ 1,84 cm.
$cV = \pm$ 1,40%.
$eM = \pm$ 0,13 cm.
$eM\% = \pm$ 0,1% (0,099 cifră rotunjită la 0,1).

Calculele pentru greutate se obțin, în mod similar, construind tabelul de distribuție (tabelul XXVIII).

Tabelul XXVIII

Greutatea la copiii de 10 ani, sex masculin, fii de muncitori din regiunea A.

Variante $= X$	Frecvențe $= n$	Deviația de la medie $= X'$	Deviația de la medie, înmulțită cu frevența $= X' \cdot n$	Deviația medie la pătrat $= (X')^2$	Deviația medie la pătrat, înmulțită cu frecvența $= (X')^2 \cdot n$
1	2	3	4	5	6
24	1	−5	5	25	25
25	4	−4	16	16	64
26	10	−3	30	9	90
27	17	−2	34	4	68
28	42	−1	42	1	42
			−127		
29	53	0	0	0	0
30	45	1	45	1	45
31	20	2	40	4	80
32	10	3	30	9	90
33	4	4	16	16	64
34	3	5	15	25	75
35	1	6	6	36	36
			+152		
	$N = 210$		$\Sigma X' \cdot n = 25$		$\Sigma (X')^2 \cdot n = 679$

Calculăm momentul 1 :

$$M_1 = \frac{\Sigma X' \cdot n}{N} = \frac{\text{coloana a 4-a}}{\text{coloana a 2-a}} = \frac{25}{210} = 0{,}12.$$

Avînd momentul 1 putem calcula media ponderată :

$$M_p = 29 + 0{,}12 = 29{,}12 \text{ kg.}$$

Pentru a putea calcula deviația standard, calculăm, în prealabil, momentul al II-lea :

$$M_2 = \frac{\Sigma (X')^2 \cdot n}{N} = \frac{\text{coloana a 6-a}}{\text{coloana a 2-a}} = \frac{679}{210} = 3{,}23.$$

Avînd momentul 1 și momentul al II-lea, putem calcula și deviația standard :

$$\sigma = \pm \sqrt{M_2 - (M_1)^2} = \pm \sqrt{3{,}23 - 0{,}12^2} = \pm \sqrt{3{,}2156} = \pm 1{,}80 \text{ cm.}$$

Coeficientul de variație va fi de 6,18% :

$$cV = \frac{\sigma \cdot 100}{Mp} = \frac{1{,}80 \cdot 100}{29{,}12} = 6{,}18\%.$$

Eroarea medie a mediei greutății va fi de ±0,12 :

$$eM = \pm \frac{\sigma}{\sqrt{N}} = \pm \frac{1{,}80}{\sqrt{210}} = \pm 0{,}12.$$

Transformată în procente, eroarea medie va fi :

$$eM\% = \frac{0{,}12 \cdot 100}{29{,}12} = 0{,}41\%.$$

Pentru aflarea deviației-tip a înălțimii și greutății de la mediile ponderate adunăm la această valoare produsul celor două momente 1 :

$$\sum n \cdot x' \cdot y' + Mx \cdot My = 591 + (0{,}1 \cdot 0{,}12) = 591{,}012.$$

Avem acum toate elementele necesare pentru calcularea coeficientului de corelație :

$$r = \frac{\Sigma n \cdot x' \cdot y' + MxMy}{N \cdot \sigma x \cdot \sigma y} = \frac{591{,}012}{210 \cdot 1{,}84 \cdot 1{,}80} = 0{,}85.$$

Cunoaștem din capitolul I al acestei lucrări că un indice de corelație de peste 0,7 denotă o corelație accentuată, strînsă. Putem, deci, afirma că între înălțimea și greutatea copiilor din eșantionul nostru există o strînsă corelație.

Pentru o mai mare exactitate calculăm și eroarea medie a coeficientului de corelație :

$$er = \pm \frac{1 - r^2}{\sqrt{N}} = \frac{1 - 0{,}72}{14{,}49} = \pm 0{,}02.$$

Luînd triplul acestei erori medii a coeficientului de corelație (0,02), putem afirma cu o probabilitate de 99,73% că acest coeficient absolut de corelație se află în intervalul dintre coeficientul de corelație calculat de noi (0,85) și triplul erorii medii a acestuia, deci între 0,79 și 0,91.

În orice caz, luînd chiar limita inferioară de 0,79, putem afirma că există, totuși, o corelație între înălțime și greutate, ținînd seama de cele arătate în capitolul I despre valoarea coeficientului de corelație (peste 0,5 putîndu-se afirma existența corelației dintre două fenomene sau însușiri).

G. INDICII DE REGRESIE ÎN DEZVOLTAREA FIZICĂ A POPULAȚIEI

Este necesar întotdeauna de a cunoaşte nu numai corelaţia simplă între două sau mai multe fenomene, ci şi mărimea cu care se modifică unul dintre fenomene, atunci cînd cel de-al doilea fenomen creşte sau scade într-o anumită măsură.

Coeficientul de regresie măsoară această modificare a două elemente corelate. În cazul nostru, coeficientul de regresie ne arată cu cît creşte sau scade greutatea în cazul în care înălțimea creşte sau scade cu o unitate de măsură, adică cu 1 cm.

Regresiunea, precum şi corelaţia pot fi în acelaşi sens, ca în cazul nostru, sau în sens opus, adică la creşterea unui element să scadă cel de-al doilea.

Calculul coeficientului de regresiune este cunoscut din capitolul I al lucrării :

$$R_{\frac{x}{y}} = r \cdot \frac{\sigma x}{\sigma y}.$$

$$R_{\frac{y}{x}} = r \cdot \frac{\sigma y}{\sigma x}.$$

Prima formulă ne arată modificarea greutății, atunci cînd înălțimea creşte sau scade cu 1 cm. Formula reprezintă, deci, regresiunea greutății față de înălțime.

Cea de a doua formulă ne arată modificarea înălțimii, atunci cînd greutatea creşte sau scade cu 1 kg. Formula reprezintă deci regresiunea înălțimii față de greutate.

Avînd elementele necesare, putem trece la calculul coeficienților de regresiune :

$$R_{\frac{x}{y}} = r \cdot \frac{\sigma x}{\sigma y} = 0{,}85 \cdot \frac{1{,}80}{1{,}84} = 0{,}83 \text{ kg};$$

$$R_{\frac{y}{x}} = r \cdot \frac{\sigma y}{\sigma x} = 0{,}85 \cdot \frac{1{,}84}{1{,}80} = 0{,}87 \text{ cm}.$$

Coeficientul de regresiune a greutății ne arată că la o creştere sau scădere cu 1 cm a înălțimii copiilor de vîrstă de 10 ani, sex masculin, fii de muncitori, greutatea creşte, respectiv scade, cu 0,83 kg. Coeficientul de regresiune a înălțimii ne arată că la o creştere sau scădere a greutății cu 1 kg a copiilor în vîrstă de 10 ani, sex masculin, fii de muncitori, înălțimea va creşte, respectiv va scădea, cu 0,87 cm.

Regresiunea ca şi corelația mai multor elemente se face combinînd două cîte două din aceste elemente în tabelele de corelație. Deci cînd corelăm 3 elemente : x, y, z, vom obține trei tabele de corelație, respectiv 3 perechi de regresiuni şi anume : x şi y, x şi z, y şi z.

ANEXA I

FORMULE MATEMATICE UTILIZATE MAI FRECVENT ÎN CERCETAREA MEDICO-SANITARĂ

Media aritmetică simplă :

$$Ma = \frac{\Sigma X}{N}$$

Ma = media aritmetică simplă ;
Σ = suma (valorilor individuale) ;
X = valorile individuale ;
N = numărul total de cazuri cercetate.

Exemplu :

$$Ma = \frac{100 + 110 + 120 + 120 + 150}{5} = 120$$

Media aritmetică ponderată :

$$Mp = \frac{\Sigma X \cdot n}{N}$$

Mp = media aritmetică ponderată ;
Σ = suma (valorilor individuale) ;
X = valorile individuale ;
n = frecvenţa (numărul de cazuri al fiecărei valori) ;
N = numărul total de cazuri cercetate.

Exemplu :

$$Mp = \frac{130 \cdot 10 + 120 \cdot 20}{10 + 20} =$$

$$= \frac{1\,300 + 2\,400}{30} = \frac{3\,700}{30} = 123$$

Momentul 1 :

$$M_1 = \frac{\Sigma X' \cdot n}{N}$$

M_1 = momentul 1 :
Σ = suma (deviaţiilor de la medie a valorilor) ;
X' = deviaţia de la medie a valorilor ;
n = frecvenţa (numărul de cazuri al fiecărei valori) ;
N = numărul total de cazuri cercetate.

Momentul al II-lea :

$$M_2 = \frac{\Sigma (X')^2 \cdot n}{N}$$

M_2 = momentul al II-lea :
Σ = suma (pătratelor deviaţiei de la medie a valorilor) ;
$(X')^2$ = pătratul deviaţiei de la medie ;
n = frecvenţa (numărul de cazuri al fiecărei valori) ;
N = numărul total de cazuri cercetate

Deviația-tip :

$$\sigma = \pm \sqrt{M_2 - (M_1)^2}$$

σ = deviația-tip (deviația standard);
M_1 = momentul 1
M_2 = momentul al II-lea.

Coeficientul de variație :

$$cV = \frac{\sigma \cdot 100}{Mp}$$

cV = coeficientul de variație;

σ = deviația-tip;
Mp = media aritmetică ponderată.

Eroarea medie :

$$eM = \frac{\sigma}{\sqrt{N}}$$

eM = eroarea medie;

σ = deviația-tip;
N = numărul total de cazuri cercetate.

Eroarea medie % :

$$eM\% = \frac{eM \cdot 100}{Mp}$$

$eM\%$ = procentul erorii medii;

eM = eroarea medie;
Mp = media aritmetică ponderată.

Eroarea probabilă :

$$ep = 0{,}6745 \cdot eM$$

ep = eroarea probabilă;
eM = eroarea medie.

Eroarea diferenței :

$$eD = \pm \sqrt{eM_1^2 + eM_2^2}$$

eD = eroarea diferenței dintre două medii;
eM = eroarea medie.

Deviația-tip în cazul frecvențelor relative :

$$\sigma = \pm \sqrt{\frac{P \cdot Q}{N}} = \pm \sqrt{\frac{P \cdot (100 - P)}{N}}$$

σ = deviația-tip;

P = probabilitatea (exprimată în procente);
Q = contraprobabilitatea (exprimată în procente);
N = numărul de cazuri cercetate.

Calculul eșantionului reprezentativ :

$$\sigma D = \pm \sqrt{\frac{P_1 \cdot (100 - P_1) + P_2 \cdot (100 - P_1)}{N}}$$

σD = diferența dispersiunii a două frecvențe relative;

$$N = \frac{P_1 \cdot (100 - P_1) + P_2 \cdot (100 - P_2)}{\sigma D^2}$$

P_1 și P_2 = probabilitatea (exprimată în procente);

N = eșantionul reprezentativ.

METODA χ^2 PENTRU MĂSURAREA CERTITUDINII DIFERENȚEI DINTRE DOUĂ SERII PE BAZA TABELULUI CVADRUPLU

	Au decedat	Nu au decedat	Total
Primul lot supus substanței A	$18_{(a_1)}$	$12_{(b_1)}$	$30_{(a_1+b_1=n_1)}$
Al doilea lot supus substanței A + B	$10_{(a_2)}$	$20_{(b_2)}$	$30_{(a_2+b_2=n_2)}$
Total	$28_{(a_1+a_2=n_3)}$	$32_{(b_1+b_2=n_4)}$	$60_{(n_1+n_2=n_3+n_4=N)}$

$$\chi^2 = \frac{(a_1 \cdot b_2 - b_1 \cdot a_2)^2}{n_1 \cdot n_2 \cdot n_3 \cdot n_4} \cdot (N-1).$$

Exemplu :

$$\chi^2 = \frac{(18 \cdot 20 - 12 \cdot 10)^2}{30 \cdot 30 \cdot 28 \cdot 32} \cdot (60 - 1) = 4{,}21.$$

Calculul stabilirii limitei superioare a probabilității în cazul evenimentelor rare :

$$m = \frac{p_0, n}{n}$$

m = indicele limitei superioare a probabilității ;

p_0 = probabilitatea de apariție în cazul în care evenimentul nu s-a produs ;

n = numărul de cazuri observate.

Coeficientul de corelație :

$$r = \frac{\Sigma n \cdot x' \cdot y'}{N \cdot \sigma x \cdot \sigma y}$$

r = coeficient de corelație ;

n = frecvența (numărul de cazuri) ;

x' = deviația de la medie a intervalelor de greutate ;

y' = deviația de la medie a intervalelor de înălțime ;

$\Sigma n \cdot x' \cdot y'$ = suma produsului deviației de la medie a greutății și înălțimii înmulțită cu frecvența ;

N = numărul total de cazuri cercetate ;

σx = deviația-tip de la media ponderată a greutății ;

σy = deviația-tip de la media ponderată a înălțimii.

Eroarea medie a coeficientului de corelație :

$$er = \frac{1 - r^2}{\sqrt{N}}$$

er = eroarea medie a coeficientului de corelație ;

r = coeficientul de corelație ;

N = numărul total de cazuri cercetate.

Coeficientul de regresiune :

$$R_{\frac{x}{y}} = r \cdot \frac{\sigma x}{\sigma y}$$

$R_{\frac{x}{y}}$ = coeficientul de regresiune al greutăţii, în raport cu înălţimea ;

$$R_{\frac{y}{x}} = r \cdot \frac{\sigma y}{\sigma x}$$

$R_{\frac{y}{x}}$ = coeficientul de regresiune al înălţimii în raport cu greutatea ;

r = coeficientul de corelaţie ;
σx = deviaţia-tip de la media ponderată a greutăţii ;
σy = deviaţia-tip de la media ponderată a înălţimii.

Ajustarea binomială a trei valori (a,b,c) :

$$X = \frac{a + 2b + c}{4}$$

X = media binomială a valorilor a, b, c.

Ajustarea binomială a patru valori (a,b,c,d) :

$$X = \frac{a + 3b + 3c + d}{8}$$

Ajustarea binomială a cinci valori (a,b,c,d,e) :

$$X = \frac{a + 4b + 6c + 4d + e}{16}$$

INDICATORI DEMOGRAFICI

Natalitatea :

$$N = \frac{\text{Numărul născuţilor vii} \cdot 1\,000}{\text{Numărul de locuitori}}$$

Fertilitatea generală :

$$Fgl = \frac{\text{Numărul născuţilor vii} \cdot 1\,000}{\text{Numărul femeilor în vîrstă de 15—49 de ani}}$$

Fertilitatea conjugală :

$$Fc = \frac{\text{Numărul născuţilor vii legitimi} \cdot 1\,000}{\text{Numărul femeilor căsătorite}}$$

Fertilitatea conjugală pe grupe de vîrstă :

$$Fc\ (20 - 24 \text{ de ani}) = \frac{\text{Numărul născuţilor vii legitimi ai mamelor în vîrstă de 20—24 de ani} \cdot 1\,000}{\text{Numărul femeilor căsătorite în vîrstă de 20—24 de ani}}$$

Fertilitatea generală după ordinea de naştere a copilului :

$$Fgl = \frac{\text{Numărul primilor născuţi vii ai femeilor în vîrstă de 15—49 ani} \cdot 1\,000}{\text{Numărul femeilor în vîrstă de 15—49 de ani}}$$

Fertilitatea conjugală după ordinea de naştere a copilului :

$$Fc = \frac{\text{Numărul primilor născuţi vii ai femeilor căsătorite în vîrstă de 15—49 de ani} \cdot 1\,000}{\text{Numărul femeilor căsătorite în vîrstă de 15—49 de ani}}$$

Mortalitatea generală :

$$Mg = \frac{\text{Numărul deceselor} \cdot 1\,000}{\text{Numărul locuitorilor}}$$

Mortalitatea pe grupe de vîrstă :

$$Mgrv = \frac{\text{Numărul deceselor de o anumită grupă de vîrstă} \cdot 1\,000}{\text{Numărul locuitorilor de aceeaşi grupă de vîrstă}}$$

Mortalitatea pe cauze :

$$M\ cauze = \frac{\text{Numărul deceselor de o anumită cauză} \cdot 100\,000}{\text{Numărul locuitorilor}}$$

Mortalitatea pe grupe de vîrstă şi cauze :

$$M\ gr\ v\ c = \frac{\text{Numărul deceselor de o anumită grupă de vîrstă şi cauză} \cdot 100\,000}{\text{Numărul locuitorilor de aceeaşi grupă de vîrstă}}$$

Mortalitatea generală corectată :

$$Mc = \frac{\Sigma Mx \cdot Sx}{1\,000\,000}$$

Mc = mortalitatea generală corectată ;
Σ = suma (mortalităţii pe ani de vîrstă înmulţită cu populaţia standard) ;
Mx = mortalitatea pe ani de vîrstă ;
Sx = mortalitatea standard pe ani de vîrstă.

Indicele vital :

$$V = \frac{\text{Numărul născuţilor vii} \cdot 100}{\text{Numărul deceselor.}} .$$

Excedent natural :

$$Ex = \frac{(\text{Numărul născuţilor vii} - \text{decese}) \cdot 1\,000}{\text{Numărul locuitorilor}}$$

Mortalitatea infantilă :

$$Mi = \frac{\text{Numărul deceselor sub 1 an} \cdot 100 \text{ (sau } 1\,000)}{\text{Numărul născuţilor vii din aceeaşi perioadă}}$$

Mortalitatea neonatală :

$$Mt = \frac{\text{Numărul deceselor sub 1 lună dintr-o anumită perioadă} \cdot 100}{\text{Numărul născuţilor vii din aceeaşi perioadă}}$$

Mortalitatea precoce :

$$Mpr = \frac{\text{Numărul deceselor sub 1 săptămînă dintr-o anumită perioadă} \cdot 100}{\text{Numărul născuţilor vii din aceeaşi perioadă}}$$

Mortinatalitatea :

$$Mn = \frac{\text{Numărul născuţilor morţi} \cdot 100}{\text{Numărul născuţilor vii}}$$

Mortalitatea infantilă corectată :

$$Mic = \left(\frac{Mi_1}{N_1} + \frac{Mi_2}{N_2}\right) \cdot 100$$

Mic = mortalitatea infantilă corectată ;
Mi_1 = decedaţi sub 1 an, născuţi în acelaşi an ;
Mi_2 = decedaţi sub 1 an, născuţi în anul anterior decesului ;
N_1 = născuţi vii din anul respectiv ;
N_2 = născuţi vii din anul anterior.

Decese la 100 de bolnavi :

$$D\% = \frac{\text{Numărul deceselor de o cauză} \cdot 100}{\text{Numărul îmbolnăvirilor de aceeaşi cauză}}$$

Letalitatea :

$$L = \frac{\text{Numărul deceselor printr-o cauză} \cdot 100}{\text{Numărul total de decese}} .$$

Decesele-mame (în primele 10 zile) :

$$Dm = \frac{\text{Numărul deceselor materne (în primele 10 zile de la naştere)} \cdot 1\,000}{\text{Numărul născuţilor (vii + morţi)}} .$$

Mortalitatea maternă :

$$Mm = \frac{\text{Numărul deceselor datorită naşterii} \cdot 1\,000}{\text{Numărul născuţilor vii}} .$$

Durata medie de viaţă :

$$ex = \frac{1}{2} + \frac{l + lx + 1 + lx + 2 + \cdots + lx + n}{lx}$$

INDICATORI DE MORBIDITATE

Morbiditatea generală :

$$Mo.\ g. = \frac{\text{Numărul cazurilor noi de îmbolnăvire} \cdot 1\,000}{\text{Numărul locuitorilor}} .$$

Morbiditatea pe grupe de vîrstă :

$$Mo\ gr\ v = \frac{\text{Numărul cazurilor noi de îmbolnăvire de o anumită grupă de vîrstă} \cdot 1000}{\text{Numărul locuitorilor de aceeaşi grupă de vîrstă}} .$$

Morbiditatea pe cauze :

$$Mo\ c = \frac{\text{Numărul cazurilor noi de îmbolnăvire de o anumită cauză} \cdot 1\,00\,000}{\text{Numărul locuitorilor}} .$$

Morbiditatea pe grupe de vîrstă şi cauze :

$$Mo\ gr\ v\ c =$$

$$= \frac{\text{Numărul cazurilor de îmbolnăvire de o anumită grupă de vîrstă şi cauză} \cdot 100\,000}{\text{Numărul locuitorilor de aceeaşi grupă de vîrstă}} .$$

Evidenţa bolnavilor :

$$E = \frac{\text{Numărul total de bolnavi (noi + vechi)} \cdot 100}{\text{Numărul locuitorilor}} .$$

Frecvența incapacității temporare de muncă:

$$F = \frac{\text{Numărul certificatelor inițiale} \cdot 100}{\text{Numărul mediu scriptic de salariați}}.$$

Gravitatea incapacității temporare de muncă:

$$G = \frac{\text{Numărul total de zile cu incapacitate temporară de muncă} \cdot 100}{\text{Numărul mediu scriptic de salariați}}.$$

Durata incapacității temporare de muncă:

$$D = \frac{\text{Numărul total de zile cu incapacitate temporară de muncă}}{\text{Numărul certificatelor inițiale}}.$$

INDICATORI DE ASISTENȚĂ MEDICO-SANITARĂ

Asigurarea populației cu personal medical:

$$As\,m = \frac{\text{Numărul locuitorilor}}{\text{Numărul medicilor}} \text{ sau } \frac{\text{Numărul medicilor} \cdot 10\,000}{\text{Numărul locuitorilor}}.$$

Asigurarea populației cu paturi de spital:

$$As\,p = \frac{\text{Numărul paturilor} \cdot 1\,000}{\text{Numărul locuitorilor}}.$$

Aglomerarea personalului medical:

$$Agl = \frac{\text{Numărul consultațiilor dintr-o perioadă}}{\text{Numărul total de ore de consultații efectiv prestate de medici. nr. medici}}.$$

Deservirea populației cu consultații:

$$Des = \frac{\text{Numărul consultațiilor}}{\text{Numărul locuitorilor}}.$$

Depistarea bolnavilor:

$$Dep = \frac{\text{Numărul zilelor de la îmbolnăvire pînă la descoperirea cazurilor}}{\text{Numărul bolnavilor depistați}}.$$

Izolarea bolnavilor:

$$Iz = \frac{\text{Numărul zilelor de la depistare pînă la izolarea cazurilor}}{\text{Numărul bolnavilor izolați}}.$$

Internarea bolnavilor:

$$In = \frac{\text{Numărul zilelor de la îmbolnăvire pînă la internarea cazurilor}}{\text{Numărul bolnavilor internați}}.$$

Utilizarea paturilor:

$$U = \frac{\text{Numărul zilelor de spitalizare (contabile)}}{\text{Numărul mediu de paturi}}.$$

Rulajul bolnavilor:

$$R = \frac{\text{Numărul bolnavilor (aflați + intrați)}}{\text{Numărul mediu de paturi}}.$$

Durata medie de spitalizare :

$$D = \frac{\text{Numărul zilelor de spitalizare (contabile)}}{\text{Numărul bolnavilor (aflaţi + intraţi)}}.$$

Durata medie de spitalizare a unei boli :

$$D\ cauză = \frac{\text{Numărul zilelor de spitalizare la bolnavii ieşiţi de o cauză}}{\text{Numărul bolnavilor ieşiţi de aceeaşi cauză}}.$$

Mortalitatea în spitale :

$$M\ sp = \frac{\text{Numărul deceselor în spital} \cdot 100}{\text{Numărul bolnavilor ieşiţi}}.$$

Decese la domiciliu :

$$D\ d = \frac{\text{Numărul deceselor la domiciliu} \cdot 100}{\text{Numărul total de decedaţi (din teritoriul respectiv)}}.$$

Decese în spital în primele 24 de ore :

$$D\ 24\ h = \frac{\text{Numărul deceselor în spital în primele 24 de ore} \cdot 100}{\text{Numărul total de decese în spital}}.$$

Mortalitatea postoperatorie :

$$D\ op = \frac{\text{Numărul deceselor postoperatorii} \cdot 100}{\text{Numărul bolnavilor operaţi}}.$$

Complicaţii postoperatorii :

$$C\ op = \frac{\text{Numărul bolnavilor cu complicaţii postoperatorii} \cdot 100.}{\text{Numărul bolnavilor operaţi}}.$$

Avorturi :

$$Av = \frac{\text{Numărul avorturilor} \cdot 100}{\text{Numărul născuţilor vii}} \text{ sau } \frac{\text{Numărul avorturilor} \cdot 1\,000}{\text{Numărul locuitorilor}}.$$

Surse de contagiune depistate

$$S\ d = \frac{\text{Numărul surselor depistate} \cdot 100}{\text{Numărul cazurilor (boli contagioase)}}.$$

$$\sigma = \pm \sqrt{\frac{\Sigma (x_1 - \bar{x})^2}{n-1}}$$

$$\sigma_m = \pm \sqrt{\frac{\Sigma (x_1 - \bar{x})^2}{n(n-1)}}$$

$$t = |\bar{x}_A - \bar{x}_B| \frac{1}{\sqrt{\sigma m_A^2 + \sigma m_B^2}}$$

test Student Fisher.

ANEXA A II-a

PĂTRATELE ȘI RĂDĂCINILE PĂTRATE ALE NUMERELOR DELA 1 LA 1500

În vederea calculării cu ușurință a unor probleme matematice mai frecvent utilizate în domeniul sanitar, și anume ridicările la pătrat și extragerea rădăcinii pătrate, redăm, în tabelele următoare, pătratele și rădăcinile pătrate ale numerelor cuprinse între 1 și 1 500.

Coloana „n" înseamnă numărul care trebuie ridicat la pătrat sau din care se extrage rădăcina pătrată, coloana „n^2" cuprinde valorile care corespund pătratelor, iar coloana „$\sqrt{n}$" rădăcinii pătrate ale numerelor din coloana „n".

Căutînd, de exemplu, pătratul și rădăcina pătrată a numărului 325, vom găsi valorile corespunzătoare rîndului 325 : 105 625 $= n^2$ și 18,0278 $= \sqrt{n}$.

Aceste tabele pot fi folosite și în cazul zecimalelor. De exemplu : 11,7 ridicat la pătrat dă 136,89. Deși în tabele nu există cifre cu zecimale, ele pot fi calculate luînd în loc de 11,7 cifra 117, ținînd seama de virgulă la rezultatul obținut. Căutînd în rîndul 117 pătratul corespunzător, găsim cifra 13 689, din care vom despărți ultimele două zecimale, adică atîtea zecimale cîte am avut în cazul nostru (11,7 . 11,7).

Dacă în loc de 11,7 vom lua cifra 11,72, ridicarea la pătrat se va face în același fel, căutînd, de data aceasta, în loc de 11,72 cifra fără zecimale (1 172) ; în rîndul corespunzător pătratului ei găsim valoarea 1 373 584. Despărțim prin virgulă ultimele patru zecimale (deoarece în cifrele 11,72. 11,72 am avut patru zecimale) și vom găsi valoarea căutată : $11{,}72^2 = 137{,}3584$.

Dacă ridicăm la pătrat 1,17 în loc de 11,7 căutăm tot valoarea corespunzătoare lui 117 = 13 689, dar în acest caz vom separa prin virgulă ultimele patru zecimale (deoarece 1,17 · 1,17 are patru zecimale), obținînd valoarea 1,3689.

În cazul rădăcinii pătrate a cifrelor zecimale, tabelul poate fi, de asemenea, utilizat. Astfel rădăcina pătrată a cifrei 11,7 trebuie căutată la 1 170 = 34,2052 și mutînd virgula la stînga cu o cifră vom găsi $\sqrt{11{,}7} = 3{,}420\ 52$.

Rădăcina pătrată din 1,17 o căutăm la 117, care ne indică valoarea 10,8167 și mutînd virgula la stînga cu o cifră vom găsi : $\sqrt{1{,}17} = 1{,}081\ 67$.

Ca o regulă de reținut pentru ușurarea calculării este și faptul că în cazul extragerii rădăcinii pătrate a cifrelor de la 1 la 99, valorile obținute vor avea la întregi o singură cifră ($\sqrt{1} = 1{,}0000$; $\sqrt{99} = 9{,}9499$) ; între 100 și 9 999 valorile obținute vor avea la întregi două cifre ($\sqrt{100} = 10{,}000\ 0$; $\sqrt{9\ 999} = 99{,}995$) ; peste 10 000 și pînă la 999 999 (inclusiv) valorile vor avea trei cifre la întregi etc.

PĂTRATELE ŞI RĂDĂCINILE PĂTRATE ALE NUMERELOR DE LA 1 LA 1 500

n	n^2	$\sqrt{n}$	n	n^2	$\sqrt{n}$
1	1	1,000 0	53	28 09	7,280 1
2	4	1,414 2	54	29 16	7,348 5
3	9	1,732 1	55	30 25	7,416 2
4	16	2,000 0	56	31 36	7,483 3
5	25	2,236 1	57	32 49	7,549 8
6	36	2,449 5	58	33 64	7,615 8
7	49	2,645 8	59	34 81	7,681 1
8	64	2,828 4			
9	81	3,000 0	60	36 00	7,746 0
			61	37 21	7,810 2
10	1 00	3,162 3	62	38 44	7,874 0
11	1 21	3,316 6	63	39 69	7,937 3
12	1 44	3,464 1	64	40 96	8,000 0
13	1 69	3,605 6	65	42 25	8,062 3
14	1 96	3,741 7	66	43 56	8,124 0
15	2 25	3,873 0	67	44 89	8,185 4
16	2 56	4,000 0	68	46 24	8,246 2
17	2 89	4,123 1	69	47 61	8,306 6
18	3 24	4,242 6			
19	3 61	4,358 9	70	49 00	8,366 6
			71	50 41	8,426 1
20	4 00	4,472 1	72	51 84	8,485 3
21	4 41	4,582 6	73	53 29	8,544 0
22	4 84	4,690 4	74	54 76	8,602 3
23	5 29	4,795 8	75	56 25	8,660 3
24	5 76	4,899 0	76	57 76	8,717 8
25	6 25	5,000 0	77	59 29	8,775 0
26	6 76	5,099 0	78	60 84	8,831 8
27	7 29	5,196 2	79	62 41	8,888 2
28	7 84	5,291 5			
29	8 41	5,385 2	80	64 00	8,944 3
			81	65 61	9,000 0
30	9 00	5,477 2	82	67 24	9,055 4
31	9 61	5,567 8	83	68 89	9,110 4
32	10 24	5,656 9	84	70 56	9,165 2
33	10 89	5,744 6	85	72 25	9,219 5
34	11 56	5,831 0	86	73 96	9,273 6
35	12 25	5,916 1	87	75 69	9,327 4
36	12 96	6,000 0	88	77 44	9,380 8
37	13 69	6,082 8	89	79 21	9,434 0
38	14 44	6,164 4			
39	15 21		90	81 00	9,486 8
		6,245 0	91	82 81	9,539 4
40	16 00	6,324 6	92	84 64	9,591 7
41	16 81	6,403 1	93	86 49	9,643 7
42	17 64	6,480 7	94	88 36	9,695 4
43	18 49	6,557 4	95	90 25	9,746 8
44	19 36	6,633 2	96	92 16	9,798 0
45	20 25	6,708 2	97	94 09	9,84 89
46	21 16	6,782 3	98	96 04	9,899 5
47	22 09	6,855 7	99	98 01	9,949 9
48	23 04	6,928 2			
49	24 01		100	1 00 00	10,000 0
		7,000 0	101	1 02 01	10,049 9
50	25 00	7,071 1	102	1 04 04	10,099 5
51	26 01	7,141 4	103	1 06 09	10,148 9
52	27 04	7,211 1	104	1 08 16	10,198 0

n	n^2	$\sqrt{n}$	n	n^2	$\sqrt{n}$
105	1 10 25	10,247 0	158	2 49 64	12,569 8
106	1 12 36	10,295 6	159	2 52 81	12,609 5
107	1 14 49	10,344 1			
108	1 16 64	10,392 3	160	2 56 00	12,649 1
109	1 18 81	10,440 3	161	2 59 21	12,688 6
			162	2 62 44	12,727 9
110	1 21 00	10,488 1	163	2 65 69	12,767 1
111	1 23 21	10,535 7	164	2 68 96	12,806 2
112	1 25 44	10,583 0	165	2 72 25	12,845 2
113	1 27 69	10,630 1	166	2 75 56	12,884 1
114	1 29 96	10,677 1	167	2 78 89	12,922 8
115	1 32 25	10,723 8	168	2 82 24	12,961 5
116	1 34 56	10,770 3	169	2 85 61	13,000 0
117	1 36 89	10,816 7			
118	1 39 24	10,862 8	170	2 89 00	13,038 4
119	1 41 61	10,908 7	171	2 92 41	13,076 7
			172	2 95 84	13,114 9
120	1 44 00	10,954 5	173	2 99 29	13,152 9
121	1 46 41	11,000 0	174	3 02 76	13,190 9
122	1 48 84	11,045 4	175	3 06 25	13,228 8
123	1 51 29	11,090 5	176	3 09 76	13,266 5
124	1 53 76	11,135 5	177	3 13 29	13,304 1
125	1 56 25	11,180 3	178	3 16 84	13 341,7
126	1 58 76	11,225 0	179	3 20 41	13,379 1
127	1 61 29	11,269 4			
128	1 63 84	11 313 7	180	3 24 00	13,416 4
129	1 66 41	11,357 8	181	3 27 61	13,453 6
			182	3 31 24	13,490 7
130	1 69 00	11,401 8	183	3 34 89	13,527 7
131	1 71 61	11,445 5	184	3 38 56	13,564 7
132	1 74 24	11,489 1	185	3 42 25	13,601 5
133	1 76 89	11,532 6	186	3 45 96	13,638 2
134	1 79 56	11,575 8	187	3 49 69	13,674 8
135	1 82 25	11,619 0	188	3 53 44	13,711 3
136	1 84 96	11,661 9	189	3 57 21	13,747 7
137	1 87 69	11,704 7			
138	1 90 44	11,747 3	190	3 61 00	13,784 0
139	1 93 21	11,789 8	191	3 64 81	13,820 3
			192	3 68 64	13,856 4
140	1 96 00	11,832 2	193	3 72 49	13,892 4
141	1 98 81	11,874 3	194	3 76 36	13,928 4
142	2 01 64	11,916 4	195	3 80 25	13,964 2
143	2 04 49	11,958 3	196	3 84 16	14,000 0
144	2 07 36	12,000 0	197	3 88 09	14,035 7
145	2 10 25	12,041 6	198	3 92 04	14,071 2
146	2 13 16	12,083 0	199	3 96 01	14,106 7
147	2 16 09	12,124 4			
148	2 19 04	12,165 5	200	4 00 00	14,142 1
149	2 22 01	12,206 6	201	4 04 01	14,177 4
			202	4 08 04	14,212 7
150	2 25 00	12,247 4	203	4 12 09	14,247 8
151	2 28 01	12,288 2	204	4 16 16	14,282 9
152	2 31 04	12,328 8	205	4 20 25	14,317 8
153	2 34 09	12,369 3	206	4 24 36	14,352 7
154	2 37 16	12,409 7	207	4 28 49	14,387 5
155	2 40 25	12,449 9	208	4 32 64	14,422 2
156	2 43 36	12,490 0	209	4 36 81	14,456 8
157	2 46 49	12,530 0			

n	n^2	$\sqrt{n}$	n	n^2	$\sqrt{n}$
210	4 41 00	14,491 4	263	6 91 69	16,217 3
211	4 45 21	14,525 8	264	6 96 96	16,248 1
212	4 49 44	14,560 2	265	7 02 25	16,278 8
213	4 53 69	14,594 5	266	7 07 56	16,309 5
214	4 57 96	14,628 7	267	7 12 89	16,340 1
215	4 62 25	14,662 9	268	7 18 24	16,370 7
216	4 66 56	14,696 9	269	7 23 61	16,401 2
217	4 70 89	14,730 9			
218	4 75 24	14,764 8	270	7 29 00	16,431 7
219	4 79 61	14,798 6	271	7 34 41	16,462 1
			272	7 39 84	16,492 4
220	4 84 00	14,832 4	273	7 45 29	16,522 7
221	4 88 41	14,866 1	274	7 50 76	16,552 9
222	4 92 84	14,899 7	275	7 56 25	16,583 1
223	4 97 29	14,933 2	276	7 61 76	16,613 2
224	5 01 76	14,966 6	277	7 67 29	16,643 3
225	5 06 25	15,000 0	278	7 72 84	16,673 3
226	5 10 76	15,033 3	279	7 78 41	16,703 3
227	5 15 29	15,066 5			
228	5 19 84	15,099 7	280	7 84 00	16,733 2
229	5 24 41	15,132 7	281	7 89 61	16,763 1
			282	7 95 24	16,792 9
230	5 29 00	15,165 8	283	8 00 89	16,822 6
231	5 33 61	15,198 7	284	8 06 56	16,852 3
232	5 38 24	15,231 5	285	8 12 25	16,881 9
233	5 42 89	15,264 3	286	8 17 96	16,911 5
234	5 47 56	15,297 1	287	8 23 69	16,941 1
235	5 52 25	15,329 7	288	8 29 44	16,970 6
236	5 56 96	15,362 3	289	8 35 21	17,000 0
237	5 61 69	15,394 8			
238	5 66 44	15,427 2	290	8 41 00	17,029 4
239	5 71 21	15,459 6	291	8 46 81	17,058 7
			292	8 52 64	17,088 0
240	5 76 00	15,491 9	293	8 58 49	17,117 2
241	5 80 81	15,524 2	294	8 64 36	17,146 4
242	5 85 64	15,556 3	295	8 70 25	17,175 6
243	5 90 49	15,588 5	296	8 76 16	17,204 7
244	5 95 36	15,620 5	297	8 82 09	17,233 7
245	6 00 25	15,652 5	298	8 88 04	17,262 7
246	6 05 16	15,684 4	299	8 94 01	17,291 6
247	6 10 09	15,716 2			
248	6 15 04	15,748 0	300	9 00 00	17,320 5
249	6 20 01	15,779 7	301	9 06 01	17,349 4
			302	9 12 04	17,378 1
250	6 25 00	15,811 4	303	9 18 09	17,406 9
251	6 30 01	15,843 0	304	9 24 16	17,435 6
252	6 35 04	15,874 5	305	9 30 25	17,464 2
253	6 40 09	15,906 0	306	9 36 36	17 492 9
254	6 45 16	15,937 4	307	9 42 49	17,521 4
255	6 50 25	15,968 7	308	9 48 64	17,549 9
256	6 55 36	16,000 0	309	9 54 81	17,578 4
257	6 60 49	16,031 2			
258	6 65 64	16,062 4	310	9 61 00	17,606 8
259	6 70 81	16,093 5	311	9 67 21	17,635 2
			312	9 73 44	17,663 5
260	6 76 00	16,124 5	313	9 79 69	17,691 8
261	6 81 21	16,155 5	314	9 85 96	17,720 0
262	6 86 44	16,186 4	315	9 92 25	17,748 2

n	n^2	$\sqrt{n}$	n	n^2	$\sqrt{n}$
316	9,98 56	17,776 4	370	13 69 00	19,235 4
317	10,04 89	17,804 5	371	13 76 41	19,261 4
318	10 11 24	17,832 6	372	13 83 84	19,287 3
319	10 17 61	17,860 6	373	13 91 29	19,313 2
			374	13 98 76	19,339 1
320	10 24 00	17,888 5	375	14 06 25	19,364 9
321	10 30 41	17,916 5	376	14 13 76	19,390 7
322	10 36 84	17,944 4	377	14 21 29	19,416 5
323	10 43 29	17,972 2	378	14 28 84	19,442 2
324	10 49 76	18,000 0	379	14 36 41	19,467 9
325	10 56 25	18,027 8			
326	10 62 76	18,055 5	380	14 44 00	19,493 6
327	10 69 29	18.083 1	381	14 51 61	19,519 2
328	10 75 84	18,110 8	382	14 59 24	19,544 8
329	10 82 41	18,138 4	383	14 66 89	19,570 4
			384	14 74 56	19,595 9
330	10 89 00	18,165 9	385	14 82 25	19,621 4
331	10 95 61	18,193 4	386	14 89 96	19,646 9
332	11 02 24	18,220 9	387	14 97 69	19,672 3
333	11 08 89	18,248 3	388	15 05 44	19,697 7
334	11 15 56	18,275 7	389	15 13 21	19,723 1
335	11 22 25	18,303 0			
336	11 28 96	18,330 3	390	15 21 00	19,748 4
337	11 35 69	18,357 6	391	15 28 81	19,773 7
338	11 42 44	18,384 8	392	15 36 64	19,799 0
339	11 49 21	18,412 0	393	15 44 49	19,824 2
			394	15 52 36	19,849 4
340	11 56 00	18,439 1	395	15 60 25	19,874 6
341	11 62 81	18,466 2	396	15 68 16	19,899 7
342	11 69 64	18,493 2	397	15 76 09	19,924 9
343	11 76 49	18,520 3	398	15 84 04	19,949 9
344	11 83 36	18,547 2	399	15 92 01	19,975 0
345	11 90 25	18,574 2			
346	11 97 16	18,601 1	400	16 00 00	20,000 0
347	12 04 09	18,627 9	401	16 08 01	20,025 0
348	12 11 04	18,654 8	402	16 16 04	20,049 9
349	12 18 01	18,681 5	403	16 24 09	20,074 9
			404	16 32 16	20,099 8
350	12 25 00	18,708 3	405	16 40 25	20,124 6
351	12 32 01	18,735 0	406	16 48 36	20,149 4
352	12 39 04	18,761 7	407	16 56 49	20,174 2
353	12 46 09	18,788 3	408	16 64 64	20,199 0
354	12 53 16	18,814 9	409	16 72 81	20,223 7
355	12 60 25	18,841 4			
356	12 67 36	18,868 0	410	16 81 00	20,248 5
357	12 74 49	18,894 4	411	16 89 21	20,273 1
358	12 81 64	18,920 9	412	16 97 44	20,297 8
359	12 88 81	18,947 3	413	17 05 69	20,322 4
			414	17 13 96	20,347 0
360	12 96 00	18,973 7	415	17 22 25	20,371 5
361	13 03 21	19,000 0	416	17 30 56	20,396 1
362	13 10 44	19,026 3	417	17 38 89	20,420 6
363	13 17 69	19,052 6	418	17 47 24	20,445 0
364	13 24 96	19,078 8	419	17 55 61	20,469 5
365	13 32 25	19,105 0			
366	13 39 56	19,131 1			
367	13 46 89	19,157 2	420	17 64 00	20,493 9
368	13 54 24	19,183 3	421	17 72 41	20,518 3
369	13 61 61	19,209 4	422	17 80 84	20,542 6

Exemple - $\sqrt{0,035} = 0,18$.

si cette $\sqrt{350} = 18,$ ...

n	n^2	$\sqrt{n}$	n	n^2	$\sqrt{n}$
423	17 89 29	20,567 0	476	22 65 76	21,817 4
424	17 97 76	20,591 3	477	22 75 29	21,840 3
425	18 06 25	20,615 5	478	22 84 84	21,863 2
426	18 14 76	20,639 8	479	22 94 41	21,886 1
427	18 23 29	20,664 0			
428	18 31 84	20,688 2	480	23 04 00	21,908 9
429	18 40 41	20,712 3	481	23 13 61	21,931 7
			482	23 23 24	21,954 5
430	18 49 00	20,736 4	483	23 32 89	21,977 3
431	18 57 61	20,760 5	484	23 42 56	22,000 0
432	18 66 24	20,784 6	485	23 52 25	22,022 7
433	18 74 89	20,808 7	486	23 61 96	22,045 4
434	18 83 56	20,832 7	487	23 71 69	22,068 1
435	18 92 25	20,856 7	488	23 81 44	22,090 7
436	19 00 96	20,880 6	489	23 91 21	22,113 3
437	19 09 69	20,904 5			
438	19 18 44	20,928 4	490	24 01 00	22,135 9
439	19 27 21	20,952 3	491	24 10 81	22,158 5
			492	24 20 64	22,181 1
440	19 36 00	20,976 2	493	24 30 49	22,203 6
441	19 44 81	21,000 0	494	24 40 36	22,226 1
442	19 53 64	21,023 8	495	24 50 25	22,248 6
443	19 62 49	21,047 6	496	24 60 16	22,271 1
444	19 71 36	21,071 3	497	24 70 09	22,293 5
445	19 80 25	21,095 0	498	24 80 04	22,315 9
446	19 89 16	21,118 7	499	24 90 01	22,338 3
447	19 98 09	21,142 4			
448	20 07 04	21,166 0	500	25 00 00	22,360 7
449	20 16 01	21,189 6	501	25 10 01	22,383 0
			502	25 20 04	22,405 4
450	20 25 00	21,213 2	503	25 30 09	22,427 7
451	20 34 01	21,236 8	504	25 40 16	22,449 9
452	20 43 04	21,260 3	505	25 50 25	22,472 2
453	20 52 09	21,283 8	506	25 60 36	22,494 4
454	20 61 16	21,307 3	507	25 70 49	22,516 7
455	20 70 25	21,330 7	508	25 80 64	22,538 9
456	20 79 36	21,354 2	509	25 90 81	22,561 0
457	20 88 49	21,377 6			
458	20 97 64	21,400 9	510	26 01 00	22,583 2
459	21 06 81	21,424 3	511	26 11 21	22,605 3
			512	26 21 44	22,627 4
460	21 16 00	21,447 6	513	26 31 69	22,649 5
461	21 25 21	21,470 9	514	26 41 96	22,671 6
462	21 34 44	21,494 2	515	26 52 25	22,693 6
463	21 43 69	21,517 4	516	26 62 56	22,715 6
464	21 52 96	21,540 7	517	26 72 89	22,737 6
465	21 62 25	21,563 9	518	26 83 24	22,759 6
466	21 71 56	21,587 0	519	26 93 61	22,781 6
467	21 80 89	21,610 2			
468	21 90 24	21,633 3	520	27 04 00	22,803 5
469	21 99 61	21,656 4	521	27 14 41	22,825 4
			522	27 24 84	22,847 3
470	22 09 00	21,679 5	523	27 35 29	22,869 2
471	22 18 41	21,702 5	524	27 45 76	22,891 0
472	22 27 84	21,725 6	525	27 56 25	22,912 9
473	22 37 29	21,748 6	526	27 66 76	22,934 7
474	22 46 76	21,771 5	527	27 77 29	22,956 5
475	22 56 25	21,794 5	528	27 87 84	22,978 3
			529	27 98 41	23,000 0

n	n^2	$\sqrt{n}$	n	n^2	$\sqrt{n}$
530	28 09 00	23,021 7	583	33 98 89	24,145 4
531	28 19 61	23,043 4	584	34 10 56	24,166 1
532	28 30 24	23,065 1	585	34 22 25	24,186 8
533	28 40 89	23,086 8	586	34 33 96	24,207 4
534	28 51 56	23,108 4	587	34 45 69	24,228 1
535	28 62 25	23,130 1	588	34 57 44	24,248 7
536	28 72 96	23,151 7	589	34 69 21	24,269 3
537	28 83 69	23,173 3			
538	28 94 44	23,194 8	590	34 81 00	24,289 9
539	29 05 21	23,216 4	591	34 92 81	24,310 5
			592	35 04 64	24,331 1
540	29 16 00	23,237 9	593	35 16 49	24,351 6
541	29 26 81	23,259 4	594	35 28 36	24,372 1
542	29 37 64	23,280 9	595	35 40 25	24,392 6
543	29 48 49	23,302 4	596	35 52 16	24,413 1
544	29 59 36	23,323 8	597	35 64 09	24,433 6
545	29 70 25	23,345 2	598	35 76 04	24,454 0
546	29 81 16	23,366 6	599	35 88 01	24,474 5
547	29 92 09	23,388 0			
548	30 03 04	23,409 4	600	36 00 00	24,494 9
549	30 14 01	23,430 7	601	36 12 01	24,515 3
			602	36 24 04	24,535 7
550	30 25 00	23,452 1	603	36 36 09	24,556 1
551	30 36 01	23,473 4	604	36 48 16	24,576 4
552	30 47 04	23,494 7	605	36 60 25	24,596 7
553	30 58 09	23,516 0	606	36 72 36	24,617 1
554	30 69 16	23,537 2	607	36 84 49	24,637 4
555	30 80 25	23,558 4	608	36 96 64	24,657 7
556	30 91 36	23,579 7	609	37 08 81	24,677 9
557	31 02 49	23,600 8			
558	31 13 64	23,622 0	610	37 21 00	24,698 2
559	31 24 81	23,643 2	611	37 33 21	24,718 4
			612	37 45 44	24,738 6
560	31 36 00	23,664 3	613	37 57 69	24,758 8
561	31 47 21	23,685 4	614	37 69 96	24,779 0
562	31 58 44	23,706 5	615	37 82 25	24,799 2
563	31 69 69	23,727 6	616	37 94 56	24,819 3
564	31 80 96	23,748 7	617	38 06 89	24,839 5
565	31 92 25	23,769 7	618	38 19 24	24,859 6
566	32 03 56	23,790 8	619	38 31 61	24,879 7
567	32 14 89	23,811 8			
568	32 26 24	23,832 8	620	38 44 00	24,899 8
569	32 37 61	23,853 7	621	38 56 41	24,919 9
			622	38 68 84	24,939 9
570	32 49 00	23,874 7	623	38 81 29	24,960 0
571	32 60 41	23,895 6	624	38 93 76	24,980 0
572	32 71 84	23,916 5	625	39 06 25	25,000 0
573	32 83 29	23,937 4	626	39 18 76	25,020 0
574	32 94 76	23,958 3	627	39 31 29	25,040 0
575	33 06 25	23,979 2	628	39 43 84	25,059 9
576	33 17 76	24,000 0	629	39 56 41	25,079 9
577	33 29 29	24,020 8			
578	33 40 84	24,041 6	630	39 69 00	25,099 8
579	33 52 41	24,062 4	631	39 81 61	25,119 7
			632	39 94 24	25,139 6
580	33 64 00	24,083 2	633	40 06 89	25,159 5
581	33 75 61	24,103 9	634	40 19 56	25,179 4
582	33 87 24	24,124 7	635	40 32 25	25,199 2

n	n^2	$\sqrt{n}$	n	n^2	$\sqrt{n}$
636	40 44 96	25,219 0	690	47 61 00	26,267 9
637	40 57 69	25,238 9	691	47 74 81	26,286 9
638	40 70 44	25,258 7	692	47 88 64	26,305 9
639	40 83 21	25,278 4	693	48 02 49	26,324 9
			694	48 16 36	26,343 9
640	40 96 00	25,298 2	695	48 30 25	26,362 9
641	41 08 81	25,318 0	696	48 44 16	26,381 8
642	41 21 64	25,337 7	697	48 58 09	26,400 8
643	41 34 49	25 357 4	698	48 72 04	26,419 7
644	41 47 36	25,377 2	699	48 86 01	26,438 6
645	41 60 25	25,396 9			
646	41 73 16	25,416 5	700	49 00 00	26,457 5
647	41 86 09	25,436 2	701	49 14 01	26,476 4
648	41 99 04	25,455 8	702	49 28 04	26,495 3
649	42 12 01	25,475 5	703	49 42 09	26,514 1
			704	49 56 16	26,533 0
650	42 25 00	25,495 1	705	49 70 25	26,551 8
651	42 38 01	25,514 7	706	49 84 36	26,570 7
652	42 51 04	25,534 3	707	49 98 49	26,589 5
653	42 64 09	25,553 9	708	50 12 64	26,608 3
654	42 77 16	25,573 4	709	50 26 81	26,627 1
655	42 90 25	25,593 0			
656	43 03 36	25 612 5	710	50 41 00	26,645 8
657	43 16 49	25,632 0	711	50 55 21	26,664 6
658	43 29 64	25,651 5	712	50 69 44	26,683 3
659	43 42 81	25,671 0	713	50 83 69	26,702 1
			714	50 97 96	26,720 8
660	43 56 00	25,690 5	715	51 12 25	26,739 5
661	43 69 21	25,709 9	716	51 26 56	26,758 2
662	43 82 44	25,729 4	717	51 40 89	26,776 9
663	43 95 69	25,748 8	718	51 55 24	26,795 5
664	44 08 96	25,768 2	719	51 69 61	26,814 2
665	44 22 25	25,787 6			
666	44 35 56	25,807 0	720	51 84 00	26,832 8
667	44 48 89	25,826 3	721	51 98 41	26,851 4
668	44 62 24	25,845 7	722	52 12 84	26,870 1
669	44 75 61	25,865 0	723	52 27 29	26,888 7
			724	52 41 76	26,907 2
670	44 89 00	25,884 4	725	52 56 25	26,925 8
671	45 02 41	25,903 7	726	52 70 76	26,944 4
672	45 15 84	25,923 0	727	52 85 29	26,962 9
673	45 29 29	25,942 2	728	52 99 84	26,981 5
674	45 42 76	25,961 5	729	53 14,41	27,000 0
675	45 56 25	25,980 8			
676	45 69 76	26,000 0	730	53 29 00	27,018 5
677	45 83 29	26,019 2	731	53 43 61	27,037 0
678	45 96 84	26,038 4	732	53 58 24	27,055 5
679	46 10 41	26,057 6	733	53 72 89	27,074 0
			734	53 87 56	27,092 4
680	46 24 00	26,076 8	735	54 02 25	27,110 9
681	46 37 61	26,096 0	736	54 16 96	27,129 3
682	46 51 24	26,115 1	737	54 31 69	27,147 7
683	46 64 89	26,134 3	738	54 46 44	27,166 2
684	46 78 56	26,153 4	739	54 61 21	27,184 6
685	46 92 25	26,172 5			
686	47 05 96	26,191 6			
687	47 19 69	26,210 7	740	54 76 00	27,202 9
688	47 33 44	26,229 8	741	54 90 81	27,221 3
689	47 47 21	26,248 8	742	55 05 64	27,239 7

n	n^2	$\sqrt{n}$	n	n^2	$\sqrt{n}$
743	55 20 49	27,258 0	796	63 36 16	28,213 5
744	55 35 36	27,276 4	797	63 52 09	28,231 2
745	55 50 25	27,294 7	798	63 68 04	28,248 9
746	55 65 16	27,313 0	799	63 84 01	28,266 6
747	55 80 09	27,331 3			
748	55 95 04	27,349 6	800	64 00 00	28,284 3
749	56 10 01	27,367 9	801	64 16 01	28,301 9
			802	64 32 04	28,319 6
750	56 25 00	27,386 1	803	64 48 09	28,337 3
751	56 40 01	27,404 4	804	64 64 16	28,354 9
752	56 55 04	27,422 6	805	64 80 25	28,372 5
753	56 70 09	27,440 8	806	64 96 36	28,390 1
754	56 85 16	27,459 1	807	65 12 49	28,407 7
755	57 00 25	27,477 3	808	65 28 64	28,425 3
756	57 15 36	27,495 5	809	65 44 81	28,442 9
757	57 30 49	27,513 6			
758	57 45 64	27,531 8	810	65 61 00	28,460 5
759	57 60 81	27,550 0	811	65 77 21	28,478 1
			812	65 93 44	28,495 6
760	57 76 00	27,568 1	813	66 09 69	28,513 2
761	57 91 21	27,586 2	814	66 25 96	28,530 7
762	58 06 44	27,604 3	815	66 42 25	28,548 2
763	58 21 69	27,622 5	816	66 58 56	28,565 7
764	58 36 96	27,640 5	817	66 74 89	28,583 2
765	58 52 25	27,658 6	818	66 91 24	28,600 7
766	58 67 56	27,676 7	819	67 07 61	28,618 2
767	58 82 89	27,694 8			
768	58 98 24	27,712 8	820	67 24 00	28,635 6
769	59 13 61	27,730 8	821	67 40 41	28,653 1
			822	67 56 84	28,670 5
770	59 29 00	27,748 9	823	67 73 29	28,688 0
771	59 44 41	27,766 9	824	67 89 76	28,705 4
772	59 59 84	27,784 9	825	68 06 25	28,722 8
773	59 75 29	27,802 9	826	68 22 76	28,740 2
774	59 90 76	27,820 9	827	68 39 29	28,757 6
775	60 06 25	27,838 8	828	68 55 84	28,775 0
776	60 21 76	27,856 8	829	68 72 41	28,792 4
777	60 37 29	27,874 7			
778	60 52 84	27,892 7	830	68 89 00	28,809 7
779	60 68 41	27,910 6	831	69 05 61	28,827 1
			832	69 22 24	28,844 4
780	60 84 00	27,928 5	833	69 38 89	28,861 7
781	60 99 61	27,946 4	834	69 55 56	28,879 1
782	61 15 24	27,964 3	835	69 72 25	28,896 4
783	61 30 89	27,982 1	836	69 88 96	28,913 7
784	61 46 56	28,000 0	837	70 05 69	28,931 0
785	61 62 25	28,017 9	838	70 22 44	28,948 2
786	61 77 96	28,035 7	839	70 39 21	28,965 5
787	61 93 69	28,053 5			
788	62 09 44	28,071 3	840	70 56 00	28,982 8
789	62 25 21	28,089 1	841	70 72 81	29,000 0
			842	70 89 64	29,017 2
790	62 41 00	28,106 9	843	71 06 49	29,034 5
791	62 56 81	28,124 7	844	71 23 36	29,051 7
792	62 72 64	28,142 5	845	71 40 25	29,068 9
793	62 88 49	28,160 3	846	71 57 16	29,086 1
794	63 04 36	28,178 0	847	71 74 09	29,103 3
795	63 20 25	28,195 7	848	71 01 04	29,120 4
			849	72 08 01	29,137 6

n	n^2	$\sqrt{n}$	n	n^2	$\sqrt{n}$
850	72 25 00	29,154 8	903	81 54 09	30,050 0
851	72 42 01	29,171 9	904	81 72 16	30,066 6
852	72 59 04	29,189 0	905	81 90 25	30,083 2
853	72 76 09	29,206 2	906	82 08 36	30,099 8
854	72 93 16	29,223 3	907	82 26 49	30,116 4
855	73 10 25	29,240 4	908	82 44 64	30,133 0
856	73 27 36	29,257 5	909	82 62 81	30,149 6
857	73 44 49	29,274 6			
858	73 61 64	29,291 6	910	82 81 00	30,166 2
859	73 78 81	29,308 7	911	82 99 21	30,182 8
			912	83 17 44	30,199 3
860	73 96 00	29,325 8	913	83 35 69	30,215 9
861	74 13 21	29,342 8	914	83 53 96	30,232 4
862	74 30 44	29,359 8	915	83 72 25	30,249 0
863	74 47 69	29,376 9	916	83 90 56	30,265 5
864	74 64 96	29,393 9	917	84 08 89	30,282 0
865	74 82 25	29,410 9	918	84 27 24	30,298 5
866	74 99 56	29,427 9	919	84 45 61	30,315 0
867	75 16 89	29,444 9			
868	75 34 24	29,461 8	920	84 64 00	30,331 5
869	75 51 61	29,478 8	921	84 82 41	30,348 0
			922	85 00 84	30,364 5
870	75 69 00	29,495 8	923	85 19 29	30,380 9
871	75 86 41	29,512 7	924	85 37 76	30,397 4
872	76 03 84	29,529 6	925	85 56 25	30,413 8
873	76 21 29	29,546 6	926	85 74 76	30,430 2
874	76 38 76	29,563 5	927	85 93 29	30,446 7
875	76 56 25	29,580 4	928	86 11 84	30,463 1
876	76 73 76	29,597 3	929	86 30 41	30,479 5
877	76 91 29	29,614 2			
878	77 08 84	29,631 1	930	86 49 00	30,495 9
879	77 26 41	29,647 9	931	86 67 61	30,512 3
			932	86 86 24	30,528 7
880	77 44 00	29,664 8	933	87 04 89	30,545 0
881	77 61 61	29,681 6	934	87 23 56	30,561 4
882	77 79 24	29,698 5	935	87 42 25	30,577 8
883	77 96 89	29,715 3	936	87 60 96	30,594 1
884	78 14 56	29,732 1	937	87 79 69	30,610 5
885	78 32 25	29,748 9	938	87 98 44	30,626 8
886	78 49 96	29,765 8	939	88 17 21	30,643 1
887	78 67 69	29,782 5			
888	78 85 44	29,799 3	940	88 36 00	30,659 4
889	79 03 21	29,816 1	941	88 54 81	30,675 7
			942	88 73 64	30,692 0
890	79 21 00	29,832 9	943	88 92 49	30,708 3
891	79 38 81	29,849 6	944	89 11 36	30,724 6
892	79 56 64	29,866 4	945	89 30 25	30,740 9
893	79 74 49	29,883 1	946	89 49 16	30,757 1
894	79 92 36	29,899 8	947	89 68 09	30,773 4
895	80 10 25	29,916 6	948	89 87 04	30,789 6
896	80 28 16	29,933 3	949	90 06 01	30,805 8
897	80 46 09	29,950 0			
898	80 64 04	29,966 6	950	90 25 00	30,822 1
899	80 82 01	29,983 3	951	90 44 01	30,838 3
			952	90 63 04	30,854 5
900	81 00 00	30,000 0	953	90 82 09	30,870 7
901	81 18 01	30,016 7	954	91 01 16	30,886 9
902	81 36 04	30,033 3	955	91 20 25	30,903 1

n	n^2	$\sqrt{n}$	n	n^2	$\sqrt{n}$
956	91 39 36	30,919 2	1 010	1 020 100	31,780 5
957	91 58 49	30,935 4	1 011	1 022 121	31,796 2
958	91 77 64	30,951 6	1 012	1 024 144	31,811 9
959	91 96 81	30,967 7	1 013	1 026 169	31,827 7
			1 014	1 028 196	31,843 4
960	92 16 00	30,983 9	1 015	1 030 225	31,859 1
961	92 35 21	31,000 0	1 016	1 032 256	31,874 8
962	92 54 44	31,016 1	1 017	1 034 289	31,890 4
963	92 73 69	31,032 2	1 018	1 035 324	31,906 1
964	92 92 96	31,048 3	1 019	1 038 361	31,921 8
965	93 12 25	31,064 4			
966	93 31 56	31,080 5	1 020	1 040 400	31,937 4
967	93 50 89	31,096 6	1 021	1 042 441	31,953 1
968	93 70 24	31,112 7	1 022	1 044 484	31,968 7
969	93 89 61	31,128 8	1 023	1 046 529	31,984 4
			1 024	1 048 576	32,000 0
970	94 09 00	31,144 8	1 025	1 050 625	32,015 6
971	94 28 41	31,160 9	1 026	1 052 676	32,031 2
972	94 47 84	31,176 9	1 027	1 054 729	32,046 8
973	94 67 29	31,192 9	1 028	1 056 784	32,062 4
974	94 86 76	31,209 0	1 029	1 058 841	32,078 0
975	95 06 26	31,225 0			
976	95 25 76	31,241 0	1 030	1 060 900	32,093 6
977	95 45 29	31,257 0	1 031	1 062 961	32,109 2
978	95 64 84	31,273 0	1 032	1 065 024	32,124 8
979	95 84 41	31,289 0	1 033	1 067 089	32,140 3
			1 034	1 069 156	32,155 9
980	96 04 00	31,305 0	1 035	1 071 225	32,171 4
981	96 23 61	31,320 9	1 036	1 073 296	32,187 0
982	96 43 24	31,336 9	1 037	1 075 369	32,202 5
983	96 62 89	31,352 8	1 038	1 077 444	32,218 0
984	96 82 56	31,368 8	1 039	1 079 521	32,233 5
985	97 02 25	31,384 7			
986	97 21 96	31,400 6	1 040	1 081 600	32,249 0
987	97 41 69	31,416 6	1 041	1 083 681	32,264 5
988	97 61 44	31,432 5	1 042	1 085 764	32,280 0
989	97 81 21	31,448 4	1 043	1 087 849	32,295 5
			1 044	1 089 936	32,311 0
990	98 01 00	31,464 3	1 045	1 092 025	32,326 5
991	98 20 81	31,480 2	1 046	1 094 116	32,341 9
992	98 40 64	31,496 0	1 047	1 096 209	32,357 4
993	98 60 49	31,511 9	1 048	1 098 304	32,372 8
994	98 80 36	31,527 8	1 049	1 100 401	32,388 3
995	99 00 25	31,543 6			
996	99 20 16	31,559 5	1 050	1 102 500	32,403 7
997	99 40 09	31,575 3	1 051	1 104 601	32,419 1
998	99 60 04	31,591 1	1 052	1 106 704	32,434 5
999	99 80 01	31,607 0	1 053	1 108 809	32,450 0
			1 054	1 110 916	32,465 4
1 000	1 000 000	31,622 8	1 055	1 113 025	32,480 8
1 001	1 002 001	31,638 6	1 056	1 115 136	32,496 2
1 002	1 004 004	31,654 4	1 057	1 117 249	32,511 5
1 003	1 006 009	31,670 2	1 058	1 119 364	32,526 9
1 004	1 008 016	31,686 0	1 059	1 121 481	32,542 3
1 005	1 010 025	31,701 7			
1 006	1 012 036	31,717 5			
1 007	1 014 049	31,733 3	1 060	1 123 600	32,557 6
1 008	1 016 064	31,749 0	1 061	1 125 721	32,573 0
1 009	1 018 081	31,764 8	1 062	1 127 844	32,588 3

n	n^2	$\sqrt{n}$	n	n^2	$\sqrt{n}$
1 063	1 129 969	32,603 7	1 116	1 245 456	33,406 6
1 064	1 132 096	32,619 0	1 117	1 247 689	33,421 6
1 065	1 134 225	32,634 3	1 118	1 249 924	33,436 5
1 066	1 136 356	32,649 7	1 119	1 252 161	33,451 5
1 067	1 138 489	32,665 0			
1 068	1 140 624	32,680 3	1 120	1 254 400	33,466 4
1 069	1 142 761	32,695 6	1 121	1 256 641	33,481 4
			1 122	1 258 884	33,496 3
1 070	1 144 900	32,710 9	1 123	1 261 129	33,511 2
1 071	1 147 041	32,726 1	1 124	1 263 376	33,526 1
1 072	1 149 184	32,741 4	1 125	1 265 625	33,541 0
1 073	1 151 329	32,756 7	1 126	1 267 876	33,555 9
1 074	1 153 476	32,771 9	1 127	1 270 129	33,570 8
1 075	1 155 625	32,787 2	1 128	1 272 384	33,585 7
1 076	1 157 776	32,802 4	1 129	1 274 641	33,600 6
1 077	1 159 929	32,817 7			
1 078	1 162 084	32,832 9	1 130	1 276 900	33,615 5
1 079	1 164 241	32,848 1	1 131	1 279 161	33,630 3
			1 132	1 281 424	33,645 2
1 080	1 166 400	32,863 4	1 133	1 283 689	33,660 1
1 081	1 168 561	32,878 6	1 134	1 285 956	33,674 9
1 082	1 170 724	32,893 8	1 135	1 288 225	33,689 8
1 083	1 172 889	32,909 0	1 136	1 290 496	33,704 6
1 084	1 175 056	32,924 2	1 137	1 292 769	33,719 4
1 085	1 177 225	32,939 3	1 138	1 295 044	33,734 2
1 086	1 179 396	32,954 5	1 139	1 297 321	33,749 1
1 087	1 181 569	32,969 7			
1 088	1 183 744	32,984 8	1 140	1 299 600	33,763 9
1 089	1 185 921	33,000 0	1 141	1 301 881	33,778 7
			1 142	1 304 164	33,793 5
1 090	1 188 100	33,015 1	1 143	1 306 449	33,808 3
1 091	1 190 281	33,030 3	1 144	1 308 736	33,823 1
1 092	1 192 464	33,045 4	1 145	1 311 025	33,837 8
1 093	1 194 649	33,060 6	1 146	1 313 316	33,852 7
1 094	1 196 836	33,075 7	1 147	1 315 609	33,867 3
1 095	1 199 025	33,090 8	1 148	1 317 904	33,882 2
1 096	1 201 216	33,105 9	1 149	1 320 201	33,896 9
1 097	1 203 409	33,121 0			
1 098	1 205 604	33,136 1	1 150	1 322 500	33,911 7
1 099	1 207 801	33,151 2	1 151	1 324 801	33,926 4
			1 152	1 327 104	33,941 1
1 100	1 210 000	33,166 2	1 153	1 329 409	33,955 8
1 101	1 212 201	33,181 3	1 154	1 331 716	33,970 6
1 102	1 214 404	33,196 4	1 155	1 334 025	33,985 3
1 103	1 216 609	33,211 5	1 156	1 336 336	34,000 0
1 104	1 218 816	33,226 5	1 157	1 338 649	34,014 6
1 105	1 221 025	33,241 5	1 158	1 340 964	34,029 4
1 106	1 223 236	33,256 6	1 159	1 343 281	34,044 0
1 107	1 225 449	33,271 6			
1 108	1 227 664	33,286 6	1 160	1 345 600	34,058 8
1 109	1 229 881	33,301 7	1 161	1 347 921	34,073 5
			1 162	1 350 244	34,088 1
1 110	1 232 100	33,316 7	1 163	1 352 569	34,102 8
1 111	1 234 321	33,331 7	1 164	1 354 896	34,117 4
1 112	1 236 544	33,346 7	1 165	1 357 225	34,132 1
1 113	1 238 769	33,361 7	1 166	1 359 556	34,146 7
1 114	1 240 906	33,376 7	1 167	1 361 889	34,161 4
1 115	1 243 225	33,391 7	1 168	1 364 224	34,176 0
			1 169	1 366 561	34,190 7

n	n^2	$\sqrt{n}$	n	n^2	$\sqrt{n}$
1 170	1 368 900	34,205 2	1 223	1 495 729	34,971 4
1 171	1 371 241	34,219 9	1 224	1 498 176	34,985 7
1 172	1 373 584	34,234 5	1 225	1 500 625	35,000 0
1 173	1 375 929	34,249 1	1 226	1 503 076	35,014 2
1 174	1 378 276	34,263 7	1 227	1 505 529	35,028 6
1 175	1 380 625	34,278 3	1 228	1 507 984	35,042 8
1 176	1 382 976	34,292 9	1 229	1 510 441	35,057 1
1 177	1 385 329	34,307 4			
1 178	1 387 684	34,322 0	1 230	1 512 900	35,071 3
1 179	1 390 041	34,336 6	1 231	1 515 361	35,085 6
			1 232	1 517 824	35,099 9
1 180	1 392 400	34,351 2	1 233	1 520 289	35,114 1
1 181	1 394 761	34,365 7	1 234	1 522 756	35,128 3
1 182	1 397 124	34,380 2	1 235	1 525 225	35,142 6
1 183	1 399 489	34,394 8	1 236	1 527 696	35,156 8
1 184	1 401 856	34,409 3	1 237	1 530 169	35,171 0
1 185	1 404 225	34,423 8	1 238	1 532 644	35,185 2
1 186	1 406 596	34,438 3	1 239	1 535 121	35,199 4
1 187	1 408 969	34,452 9			
1 188	1 411 344	34,467 4	1 240	1 537 600	35,213 7
1 189	1 413 721	34,481 9	1 241	1 540 081	35,227 8
			1 242	1 542 564	35,242 0
1 190	1 416 100	34,496 4	1 243	1 545 049	35,256 2
1 191	1 418 481	34,510 8	1 244	1 547 536	35,270 4
1 192	1 420 864	34,525 3	1 245	1 550 025	35,284 5
1 193	1 423 249	34,539 8	1 246	1 552 516	35,298 7
1 194	1 425 636	34,554 3	1 247	1 555 009	35,312 8
1 195	1 428 025	34,568 8	1 248	1 557 504	35,327 0
1 196	1 430 416	34,583 2	1 249	1 560 001	35,341 2
1 197	1 432 809	34,597 7			
1 198	1 435 204	34,612 1	1 250	1 562 500	35,355 3
1 199	1 437 601	34,626 6	1 251	1 565 001	35,369 5
			1 252	1 567 504	35,383 6
1 200	1 440 000	34,641 0	1 253	1 570 009	35,397 7
1 201	1 442 401	34,655 5	1 254	1 572 516	35,411 9
1 202	1 444 804	34,669 8	1 255	1 575 025	35,426 0
1 203	1 447 209	34,684 3	1 256	1 577 536	35,440 1
1 204	1 449 616	34,698 7	1 257	1 580 049	25,454 2
1 205	1 452 025	34,713 1	1 258	1 582 564	35,468 3
1 206	1 454 436	34,727 5	1 259	1 585 081	35,482 4
1 207	1 456 849	34,741 9			
1 208	1 459 264	34,756 3	1 260	1 587 600	35,496 5
1 209	1 461 681	34,770 7	1 261	1 590 121	35,510 5
			1 262	1 592 644	35,524 6
1 210	1 464 100	34,785 0	1 263	1 595 169	35,538 7
1 211	1 466 521	34,799 4	1 264	1 597 696	35,552 8
1 212	1 468 944	34,813 8	1 265	1 600 225	35,566 9
1 213	1 471 369	34,828 1	1 266	1 602 756	35,580 9
1 214	1 473 796	34,842 5	1 267	1 605 289	35,595 0
1 215	1 476 225	34,856 8	1 268	1 607 824	35,609 0
1 216	1 478 656	34,871 2	1 269	1 610 361	35,623 1
1 217	1 481 089	34,885 5			
1 218	1 483 524	34,899 9	1 270	1 612 900	35,637 1
1 219	1 485 961	34,914 2	1 271	1 615 441	35,651 1
			1 272	1 617 984	35,665 1
1 220	1 488 400	34,928 5	1 273	1 620 529	35,679 1
1 221	1 490 841	34,942 8	1 274	1 623 076	35,693 1
1 222	1 493 284	34,957 1	1 275	1 625 625	35,707 2

n	n^2	$\sqrt{n}$	n	n^2	$\sqrt{n}$
1 276	1 628 176	35,721 1	1 330	1 768 900	36,469 2
1 277	1 630 729	35,735 1	1 331	1 771 561	36,482 8
1 278	1 633 284	35,749 1	1 332	1 774 224	36,496 6
1 279	1 635 841	35,763 1	1 333	1 776 889	36,510 3
			1 334	1 779 556	36,524 0
1 280	1 638 400	35,777 1	1 335	1 782 225	36,537 6
1 281	1 640 961	35,791 0	1 336	1 784 896	36,551 3
1 282	1 643 524	35,805 0	1 337	1 787 569	36,565 0
1 283	1 646 089	35,819 0	1 338	1 790 244	36,578 7
1 284	1 648 656	35,833 0	1 339	1 792 921	36,592 4
1 285	1 651 225	35,846 9			
1 286	1 653 796	35,860 9	1 340	1 795 600	36,606 0
1 287	1 656 469	35,874 8	1 341	1 798 281	36,619 7
1 288	1 658 944	35,888 8	1 342	1 800 964	36,633 4
1 289	1 661 521	35,902 6	1 343	1 803 649	36,647 0
			1 344	1 806 336	36,660 6
1 290	1 664 100	35,916 6	1 345	1 809 025	36,674 2
1 291	1 666 681	35,930 5	1 346	1 811 716	36,687 9
1 292	1 669 264	35,944 5	1 347	1 814 409	36,701 5
1 293	1 671 849	35,958 3	1 348	1 817 104	36,715 1
1 294	1 674 436	35,972 2	1 349	1 819 801	36,728 7
1 295	1 677 025	35,986 1			
1 296	1 679 616	36,000 0	1 350	1 822 500	36,742 3
1 297	1 682 209	36,013 9	1 351	1 825 201	36,756 0
1 298	1 684 804	35,027 8	1 352	1 827 904	36,769 5
1 299	1 687 401	36,041 6	1 353	1 830 609	36,783 2
			1 354	1 833 316	36,796 8
1 300	1 690 000	36,055 5	1 355	1 836 025	36,810 3
1 301	1 692 601	36,069 3	1 356	1 838 736	36,823 9
1 302	1 695 204	36,083 2	1 357	1 841 449	36,837 5
1 303	1 697 809	36,097 1	1 358	1 844 164	36,851 1
1 304	1 700 416	36,110 9	1 359	1 846 881	36,864 6
1 305	1 703 025	36,124 8			
1 306	1 705 636	36,138 6	1 360	1 849 600	36,878 2
1 307	1 708 249	36,152 5	1 361	1 852 321	36,891 7
1 308	1 710 864	36,166 3	1 362	1 855 044	36,905 3
1 309	1 713 481	36,180 1	1 363	1 857 769	36,918 8
			1 364	1 860 496	36,932 3
1 310	1 716 100	36,193 9	1 365	1 863 225	36,946 0
1 311	1 718 721	36,207 7	1 366	1 865 956	36,959 5
1 312	1 721 344	36,221 6	1 367	1 868 689	36,973 0
1 313	1 723 969	36,235 4	1 368	1 871 424	36,986 5
1 314	1 726 596	36,249 1	1 369	1 874 161	37,000 0
1 315	1 729 225	36,262 9			
1 316	1 731 856	36,276 7	1 370	1 876 900	37,013 5
1 317	1 734 489	36,290 5	1 371	1 879 641	37,027 0
1 318	1 737 124	36,304 3	1 372	1 882 384	37,040 5
1 319	1 739 761	36,318 0	1 373	1 885 129	37,054 0
			1 374	1 887 876	37,067 5
1 320	1 742 400	36,331 8	1 375	1 890 625	37,081 0
1 321	1 745 041	36,345 5	1 376	1 893 376	37,094 5
1 322	1 747 684	36,359 3	1 377	1 896 129	37,107 9
1 323	1 750 329	36,373 1	1 378	1 898 884	37,121 4
1 324	1 752 976	36,386 8	1 379	1 901 641	37,134 9
1 325	1 755 625	36,400 6			
1 326	1 758 276	36,414 3			
1 327	1 760 929	36,428 0	1 380	1 904 400	37,148 4
1 328	1 763 584	36,441 7	1 381	1 907 161	37,161 8
1 329	1 766 241	36,455 5	1 382	1 909 924	37,175 3

n	n^2	$\sqrt{n}$	n	n^2	$\sqrt{n}$
1 383	1 912 689	37,188 7	1 436	2 062 096	37,894 6
1 384	1 915 456	37,202 2	1 437	2 064 969	37,907 8
1 385	1 918 225	37,215 6	1 438	2 067 844	37,921 0
1 386	1 920 996	37,229 0	1 439	2 070 721	37,934 1
1 387	1 923 769	37,242 4			
1 388	1 926 544	37,255 9	1 440	2 073 600	37,947 3
1 389	1 929 321	37,269 3	1 441	2 076 481	37,960 5
			1 442	2 079 364	37,973 7
1 390	1 932 100	37,282 7	1 443	2 082 249	37,986 8
1 391	1 934 881	37,296 1	1 444	2 085 136	38,000 0
1 392	1 937 664	37,309 5	1 445	2 088 025	38,013 2
1 393	1 940 449	37,322 9	1 446	2 090 916	38,026 3
1 394	1 943 236	37,336 3	1 447	2 093 809	38,039 5
1 395	1 946 025	37,349 7	1 448	2 096 704	38,052 6
1 396	1 948 816	37,363 0	1 449	2 099 601	38,065 7
1 397	1 951 609	37,376 5			
1 398	1 954 404	37,389 9	1 450	2 102 500	38,078 8
1 399	1 957 201	37,403 2	1 451	2 105 401	38,092 0
			1 452	2 108 304	38,105 1
1 400	1 960 000	37,416 6	1 453	2 111 209	38,118 3
1 401	1 962 801	37,429 9	1 454	2 114 116	38,131 3
1 402	1 965 604	37,443 3	1 455	2 117 025	38,144 5
1 403	1 968 409	37,456 7	1 456	2 119 936	38,157 5
1 404	1 971 216	37,470 0	1 457	2 122 849	38,170 7
1 405	1 974 025	37,483 3	1 458	2 125 764	38,183 8
1 406	1 976 836	37,496 6	1 459	2 128 681	38,196 9
1 407	1 979 649	37,510 0			
1 408	1 982 464	37,523 3	1 460	2 130 600	38,210 0
1 409	1 985 281	37,536 6	1 461	2 134 521	38,223 0
			1 462	2 137 444	38,236 1
1 410	1 988 100	37,550 0	1 463	2 140 369	38,249 2
1 411	1 990 921	37,563 3	1 464	2 143 296	38,262 2
1 412	1 993 744	37,576 6	1 465	2 146 225	38,275 3
1 413	1 996 569	37,589 9	1 466	2 149 156	38,288 4
1 414	1 999 396	37,603 2	1 467	2 152 089	38,301 5
1 415	2 002 225	37,616 5	1 468	2 155 024	38,314 5
1 416	2 005 056	37,629 8	1 469	2 157 961	38,327 5
1 417	2 007 889	37,643 1			
1 418	2 010 724	37,656 3	1 470	2 160 900	38,340 6
1 419	2 013 561	37,669 6	1 471	2 163 841	38,353 6
			1 472	2 166 784	38,366 6
1 420	2 016 400	37,682 9	1 473	2 169 729	38,379 7
1 421	2 019 241	37,696 2	1 474	2 172 676	38,392 7
1 422	2 022 284	37,709 4	1 475	2 175 625	38,405 7
1 423	2 024 929	37,722 7	1 476	2 178 576	38,418 7
1 424	2 027 776	37,735 9	1 477	2 181 529	38,431 7
1 425	2 030 625	37,749 2	1 478	2 184 484	38,444 7
1 426	2 033 476	37,762 5	1 479	2 187 441	38,457 8
1 427	2 036 329	37,775 7			
1 428	2 039 184	37,788 9	1 480	2 190 400	38,470 8
1 429	2 042 041	37,802 1	1 481	2 193 361	38,483 8
			1 482	2 196 324	38,496 7
1 430	2 044 900	37,815 3	1 483	2 199 289	38,509 7
1 431	2 047 761	37,828 6	1 484	2 202 256	38,522 7
1 432	2 050 624	37,841 8	1 485	2 205 225	38,535 7
1 433	2 053 489	37,855 0	1 486	2 208 196	38,548 7
1 434	2 056 356	37,868 2	1 487	2 211 169	38,561 6
1 435	2 059 225	37,881 4			

n	n^2	$\sqrt{n}$	n	n^2	$\sqrt{n}$
1 488	2 214 144	38,574 6	1 494	2 232 036	38,652 3
1 489	2 217 121	38,587 6	1 495	2 235 025	38,665 2
			1 496	2 238 016	38,678 2
1 490	2 220 100	38,600 5	1 497	2 241 009	38,691 1
1 491	2 223 081	38,613 5	1 498	2 244 004	38,704 0
1 492	2 226 064	38,626 4	1 499	2 247 001	38,716 9
1 493	2 229 049	38,639 4	1 500	2 250 000	38,729 9

$\bar{x}$ = suma.

Splină - 1965.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{1d}$$

$$1^2 d = \frac{\sum (x - \bar{x}_1)^2 + \sum (x - \bar{x}_2)}{N_1 N_2 (N_1 + N_2 - 2)} \cdot \frac{2 N_1 + N_2}{}$$

$$1d = \sqrt{1^2 d}$$

$\sum (x - \bar{x}_1)^2$ = suma pătratelor diferențelor dintre termenii 1 de comparat

$\sum (x - \bar{x}_2)$ suma pătratelor diferențelor dintre termenii 2 de comparat

$2(N_1 + N_2)$

BIBLIOGRAFIE

VIGDORCIC N. A. — Aplicarea statisticii în clinică, Institutul de Igienă şi Boli Profesionale Leningrad, 1945.

DARMOIS G. — Statistique et applications , Ed. A. Colin, Paris, 1946.

MORICE E., TISSERAND M., REBOUL J. — Méthodes statistiques en médecine et en biologie, Ed. Masson, Paris, 1947.

HOSEMANN H. — Die Grundlagen der statistischen Methoden für Mediziner und Biologen, Ed. G. Thieme, Suttgart, 1949.

VANDERVAEL F. — Biometrie humaine, Ed. Masson, Paris.

GRACIOV N. G. — Statisticeskie grupirovki, Gostatizdat, Moscova, 1951.

* * * — Teoria statistiki, Gostatizdat, Moscova, 1953.

MARTINI P. — Methodenlehre der therapeutisch-klinischen Forschung, Göttingen, Heidelberg, Berlin, 1953.

BOIARSKI A. I. — Metode statistice în cercetările medicale experimentale, Medghiz, Moscova, 1955.

KUVŞINIKOV P.A. — Metoda statistică în cercetările clinice, Medghiz, Moscova, 1955 .

MILLS F. C. — Statistical methods, ed. a III-a, Columbia University New York, 1955.

SĂHLEANU V. — Metode matematice de cercetare medico-biologică, Ed. medicală, Bucureşti, 1957.

LAMOTTE M. — Initiation aux methodes statistiques en biologie, Ed. Masson, Paris, 1957.

SWAROOP S. — Statistical considerations and methodology in malaria eradication, Organisation Mondiale de la Santé, Geneva, 1959.

LISON L. — Statistique appliquée à la biologie expérimentale, Ed. Gauthier-Villars, Paris, 1958.

BESMERTNÎI B. S., AKATOV A. K. — *Ž. Microbiol. (Mosk.).*, 1959, nr. 5.

SEREBREANNIKOV V. D. — *Radiologia medicală*, 1960, nr. 3.

MORISSON R. T. — *Canad. med. Ass.*, 1961, vol. 84.

REISENAUER R. — *Čas, Lék. čes.*, 1961, nr. 17.

STEINBACH M. — Prelucrarea statistică în medicină şi biologie, Ed. Acad. R.P.R., Bucureşti, 1961.

G. L'ELTORE Epidémiologie statistique, ed. Richerche, Roma.

TABLA DE MATERII

CAPITOLUL I

NOŢIUNI GENERALE

1. Introducere 5
2. Metodologia cercetării 7
3. Legea numerelor mari 11
4. Calculul probabilităţilor 15
5. Valoarea medie 19
6. Metode simplificate pentru calcularea valorii medii 25
7. Dispersiunea de la medie 30
8. Eroarea medie a mediei aritmetice. 38
9. Măsurarea dispersiunii şi erorii medii a frecvenţelor relative. Semnificaţia diferenţei dintre frecvenţele relative 42
10. Măsurarea erorii medii a mediei aritmetice în cazul frecvenţelor mici. 43
11. Măsurarea dispersiunii în cazul evenimentelor rare „Legea numerelor mici”. . . . 44
12. Mărimea eşantionului reprezentativ 46
13. Măsurarea diferenţei dintre seriile de variaţie prin metoda χ^2 (chi pătrat). . . . 48
14. Corelaţia dintre fenomene. Regresiunea 49

CAPITOLUL AL II-LEA

METODE MATEMATICE ÎN CERCETĂRILE MEDICALE EXPERIMENTALE

1. Introducere 57
2. Lotul experimental şi lotul martor 59
3. Rezultatul mediu, dispersiunea, certitudinea diferenţei a două medii. Calculul eşantionului reprezentativ 61
4. Calculul certitudinii diferenţei dintre doi indici prin metoda χ^2 (chi pătrat). 65
5. Calculul dispersiunii cu ajutorul tabelelor 67
6. Calculul dispersiunii în cazul evenimentelor rare 67
7. Calculul eşantionului reprezentativ 69

CAPITOLUL AL III-LEA

METODE MATEMATICE ÎN CERCETĂRILE CLINICE

1. Noţiuni generale 72
2. Calculul valorii medii, al gradului de omogenitate şi variabilitate. Eroarea medie a valorii medii. Certitudinea diferenţei dintre două sau mai multe rezultate 75
3. Calculul dispersiunii în cazul evenimentelor rare 77
4. Calculul diferenţei în cazul frecvenţelor mici 78
5. Calculul eşantionului reprezentativ în observaţiile clinice 80

CAPITOLUL AL IV-LEA

METODE MATEMATICE ÎN CERCETĂRILE STĂRII DE SĂNĂTATE A POPULAŢIEI

1. Metode de cercetare demografică. Partea generală 82
2. Metode de cercetare demografică. Partea specială 91
   A. Metode de cercetare privind numărul și structura populației 91
   B. Metode de cercetare privind dinamica populației 95
   C. Metode de apreciere a omogenității eșantionului și a gradului de variabilitate a fenomenelor demografice 99
   D. Metode de apreciere a certitudinii diferenței dintre frecvențele relative în demografie 100
   E. Metode de calcul a eșantionului reprezentativ în demografie 101
   F. Metode de corectare a indicilor demografici 103
      a) Metode de corectare a indicelui de natalitate 104
      b) Metode de corectare a indicelui de mortalitate generală 105
      c) Metode de corectare a indicelui de mortalitate infantilă 110
   G. Durata medie de viață 113
   H. Metode de stabilire a tendinței fenomenelor demografice 116
3. Metode de selecție în studiul morbidității 119
   A. Numărul eșantioanelor 119
   B. Mărimea eșantioanelor 120
   C. Metoda de apreciere a omogenității eșantionului și a gradului ei de variabilitate în studiul morbidității 121
   D. Metoda de calcul a certitudinii diferenței dintre frecvențele morbidității a două eșantioane 122
   E. Metoda de calcul a eșantionului reprezentativ 123
   F. Metoda de corectare a morbidității 124
   G. Tendințe în dinamica morbidității 124
4. Cercetarea de selecție în domeniul dezvoltării fizice a populației. 126
   A. Numărul eșantionului 127
   B. Mărimea eșantionului 128
   C. Calculul valorii medii a eșantionului 129
   D. Calculul dispersiunii de la medie 130
   E. Calculul erorii medii 131
   F. Corelația în domeniul dezvoltării fizice a populației 132
   G. Indicii de regresie în dezvoltarea fizică a populației 135

ANEXA I

Formule matematice utilizate mai frecvent în cercetarea medico-sanitară. . . 137

ANEXA A II-a

Pătratele și rădăcinile pătrate ale numerelor de la 1 la 1 500. 145

- (EROAREA MEDIEI) -

$$\sigma_m = \pm \sqrt{\frac{\sum (\bar{x} - x)^2}{n(n-1)}}$$

$\sum (\bar{x} - x)^2$ = suma pătratelor abaterilor.

n = nr. de cazuri.

Distribuția lui $\bar{T}(P)$

$$s^2_d = \left[\sum (x - \bar{x}_1)^2 + \sum (x - \bar{x}_2)^2 \right] \cdot \frac{N_1 + N_2}{N_1 \cdot N_2 (N_1 + N_2 - 2)}$$

$$s_d = \sqrt{s^2_d} \qquad t = \frac{\bar{x}_2 - \bar{x}_1}{s_d}$$

$\sum (x - x_1)^2$ = suma pătratelor abaterilor cazurilor din I lot

$\sum (x - x_2)^2$ = suma pătratelor def. cazuri din lot II

N_1 = nr. cazuri din I lot; N_2 = nr. cazuri din lot II

s = suma pătratelor diferențelor)

×

Redactor responsabil : Farm. Emil Gheorghiu
Tehnoredactor : Trîmbițoni Ruth

Dat la cules : 18.05.1963. Bun de tipar : 27.07.1963. Apărut : 1963. Tiraj : 3.000 + 120 ex. broșate. Hîrtie : velină ilustrații de 80 g/m², 700×1000/16. Coli editoriale : 10,45. Coli de tipar : 10,25. A. : 5245/1963. C. Z. pentru bibliotecile mari : 51.616. C. Z. pentru bibliotecile mici : 51.616.

I. P. I. str. Brezoianu nr. 23—25 București R.P.R., comanda 2438.

64

- (EROAREA MEDIEI) -

$$\sigma_m = \pm \sqrt{\frac{\sum (\bar{x} - x)^2}{n(n-1)}}$$

$\sum (\bar{x} - x)^2$ = suma pătratelor abaterilor.

n = nr. de cazuri.

Distribuția lui $\bar{x}$ (P)

$$s^2_d = \left[\sum (x - \bar{x}_1)^2 + \sum (x - \bar{x}_2)^2 \right] \cdot \frac{N_1 + N_2}{N_1 \cdot N_2 (N_1 + N_2 - 2)}$$

$$s_d = \sqrt{s^2_d} \qquad t = \frac{\bar{x}_2 - \bar{x}_1}{s_d}$$

$\sum (x - x_1)^2$ = suma pătratelor abaterilor cazurilor lui $\bar{x}_1$

$\sum (x - x_2)^2$ = suma pătratelor abaterilor cazurilor lui $\bar{x}_2$

N_1 = nr. cazuri lui $\bar{x}_1$; N_2 = nr. cazuri lui $\bar{x}_2$

s = suma pătratelor diferențelor

Redactor responsabil : Farm. Emil Gheorghiu
Tehnoredactor : Trimbițoni Ruth

Dat la cules : 18.05.1963. Bun de tipar : 27.07.1963. Apărut : 1963. Tiraj : 3.000 + 120 ex. broșate. Hîrtie : velină ilustrații de 80 g/m², 700×1000/16. Coli editoriale : 10,45. Coli de tipar : 10,25. A. : 5245/1963. C. Z. pentru bibliotecile mari : 51.616. C. Z. pentru bibliotecile mici : 51.616.

I. P. I. str. Brezoianu nr. 23—25 București R.P.R.,
comanda 2438.

A. Balan.

$\sqrt{24,11} = 4,90$

$\sqrt{24} = 4,89$ } 0,11
$\sqrt{25} = 5,00$ }

24,52

$\sqrt{24}$ — 4,89 $^{+0,01}$ }
$\sqrt{25}$ 5,00 } 0,11

1 — 0,11

0,52 — x

$x = 0,52 \cdot 0,11$

$t = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\Sigma(\bar{x}_1 - \bar{x}_2)^2 + \Sigma(\bar{x}_1 - \bar{x}_2)^2 \cdot \dfrac{N_1 + N_2}{N_1 \cdot N_2 (N_1 + N_2 - 2)}}}$

$t = (\bar{x}_1 - \bar{x}_2) \sqrt{\dfrac{N_1 \cdot N_2}{\Sigma(\bar{x}_1 - \bar{x}_2^2) + \Sigma(x_1 - \bar{x}_2)^2 \times (N_1 + N_2)}{N_1 + N_2 - 2}}$.

$\lambda^2 d = \left(S x_1^2 + S x_2^2\right) \dfrac{N_1 + N_2}{N_1 \cdot N_2 (N_1 + N_2 - 2)}$

$t = \dfrac{\bar{x} - x_2}{\lambda d}$

em sau $\sigma_M = \pm \sqrt{\frac{\Sigma (x_1 - \bar{x})^2}{n(n-1)}} = \frac{\sigma}{\sqrt{n}}$

Σ = suma
em = eroare medie = sau σ_m
σ = deviatia standard
x_1 = valori individuale
$\bar{x}$ = valoarea medie
n = numar. cazuri

$t = (\bar{x}_A - \bar{x}_B) \frac{1}{\sqrt{em_A^2 + em_B^2}}$ → p. diferenta egal.
sau $\sigma m_A^2 + \sigma m_B^2$)

P = din tabel pg 315. Distributia lui T = se cauta valorile din cele 2 coloane de comparat, din total se vede 2 (nr. categorii). Cifra obtinuta se ia din tabel - din margine, pe linia pe care se merge orizontal pana la o valoare apropiata de t - obtinut din calcul. De la ea. cifra se urca pe coloana o cifra sus P%, adica 0,1 ; 1, 2, 5 etc.
1 = P = 0,01 5 = P = 0,05.

Ex anim
lotul I - 6 6+16=22-2.
" II - 16
" III - 8

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\sum (\bar{x}_1 - x_M)^2 + \sum (\bar{x} - x_M)^2 \frac{N_1 + N_2}{N_1 \cdot N_2 (N_1 + N_2 - 2)}}}$$

Se folosesc în cazul în care se calculează după următ. formulă

$$\bar{T} = (\bar{x}_A - \bar{x}_B) \sqrt{\frac{N_A \cdot N_B}{s^2 (N_A + N_B}}$$

$$s^2 = \frac{N_A \cdot (N_A - 1)\, s^2 N_A + N_B (N_B - 1)\, s m^2 B)}{N_A + N_B - 2}$$

$$s^2 = \frac{\Sigma_1 (x_1 - \bar{x}_2)^2 + \Sigma_2 (x_2 - \bar{x}_1)^2}{n_1 + n_2 - 2}$$

$$s^2 d = (s x_1^2 + s x_2^2) \frac{N_1 + N_2}{N_1 N_2 (N_1 + N_2 - 2)}$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s d} \qquad s d = \sqrt{s^2 d}.$$

Eroarea diferenței a două medii aritmetice es (eroare diferenței dintre medii) =

$$e\Delta = \pm \sqrt{e M_1^2 + e M_2^2}$$

Dacă eroare medie a diferenței între cele două medii aritmetice paralele (es) se află egală cu rădăcina pătrată a sumei celor două erori medii ridicate la pătrat

Ex. 50,64 - 50,33 = 0,31. Răd. pătrată din suma pătratelor erorii medii se fr -

$$e\Delta = \pm \sqrt{e M_1^2 + e M_2^2} = \pm \sqrt{0,1^2 + 0,08^2} = \pm \sqrt{0,0164}$$

$$\pm 0,128.$$

eroare diferenței

Metode matematice în clinică,
laboratoare și ocrotirea sănătății

M. Balan.

$p <= max$

$p >= max$

$> p = m. max$

$< p = m. min$

$$\sigma = \pm \sqrt{\frac{\sum (x_1 - \bar{x})^2}{n-1}}$$

$$\sigma_m = \pm \sqrt{\frac{\sum (x_1 - \bar{x})^2}{n(n-1)}}$$

σ_m = eroare medie

σ = abaterea standard

x_1 = valorile individuale

$\bar{x}$ = valoarea medie

n = numărul cazurilor

x_1 = val individ

$\bar{x}$ = Testul Student. Findcu suma patratelor abaterii

$$em = \frac{\sigma}{\sqrt{n}}$$

$$t = (\bar{x}_A - \bar{x}_B) \frac{1}{\sqrt{\sigma m_A^2 + \sigma m_B^2}}$$

Distribuția lui T: se adună animalele din cele două loturi, din total se scade 3 → (numărul loturilor) cifra obținută în dreptul coloanei din margine orizontal până la coloana specială de t obținut din calcul - de la cifrele ne ridicăm pe coloane până la rândul loturi inegale -

$$s^2 = \frac{\sum_1 (x_1 - \bar{x}_1)^2 + \sum (x_2 - \bar{x}_2)^2}{n_1 + n_2 - 2}$$

$$t = (\bar{x}_A - \bar{x}_B) \sqrt{\frac{n_A \cdot n_B}{s^2 (n_A + n_B)}}$$

$\sum$ = suma ~~abaterilor~~ la 2 / suma patratelor abaterii